A generalization of Schur-Weyl duality with applications in quantum estimation

Iman Marvian[1,2] and Robert W. Spekkens[1]

[1]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada N2L 2Y5*
[2]*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

(Dated: December 1, 2011)

Schur-Weyl duality is a powerful tool in representation theory which has many applications to quantum information theory. We provide a generalization of this duality and demonstrate some of its applications. In particular, we use it to develop a general framework for the study of a family of quantum estimation problems wherein one is given n copies of an unknown quantum state according to some prior and the goal is to estimate certain parameters of the given state. In particular, we are interested to know whether collective measurements are useful and if so to find an upper bound on the amount of entanglement which is required to achieve the optimal estimation. In the case of pure states, we show that commutativity of the set of observables that define the estimation problem implies the sufficiency of unentangled measurements.

PACS numbers:

Contents

I. INTRODUCTION

Schur-Weyl duality is a duality between two subgroups of the general linear group on $(\mathbb{C}^d)^{\otimes n}$: the collective action of the unitary group $\mathrm{U}(d)$, and the canonical representation of the group $\mathcal{S}_n$ of permutations of the n systems. It asserts that there is a one-to-one map between the irreducible representations of the two groups, and that their product is multiplicity-free. Alternatively, one can characterize the duality as the fact that the complex algebra spanned by one of these groups is the commutant of the one spanned by the other. The generalization we derive here is also between two subgroups of the general linear group on $(\mathbb{C}^d)^{\otimes n}$. One is the collective action of a subgroup G of $\mathrm{U}(d)$ that has a particular property, namely, that it is equal to the centralizer of its centralizer in $\mathrm{U}(d)$. We call such a group a *gauge* group (for reasons that will be explained shortly). The other is the group closure of the local action of G' (the centralizer of G) and the canonical representation of the permutation

group $\mathcal{S}_n$. Schur-Weyl duality is included as the special case where $G = \mathrm{U}(d)$.

Just as Schur-Weyl duality has many applications to quantum information theory and quantum algorithms (see [1] and [2] for a review), so too does this generalization. This article will explore some of these applications.

One such application is to finding noiseless subsystems (this is considered in Sec. IV A 1). However, most of the applications will rely on a particular consequence which connects global symmetries with local symmetries, considered in Sec. IV B.

For M an arbitrary operator on $(\mathbb{C}^d)^{\otimes n}$, we say that M has *global symmetry* with respect to the subgroup H of $\mathrm{U}(d)$ if it is invariant under the collective action of H, i.e.,

$$\forall V \in H : \; V^{\otimes n} M V^{\dagger \otimes n} = M, \tag{1.1}$$

and we say that M has *local symmetry* with respect to H if it is invariant under the local action of H, i.e.,

$$\begin{aligned} &\forall V \in H \;\text{ and }\; \forall k : 0 \leq k \leq n-1, \\ &(I^{\otimes k} \otimes V \otimes I^{\otimes(n-k-1)}) M (I^{\otimes k} \otimes V^\dagger \otimes I^{\otimes(n-k-1)}) = M \end{aligned} \tag{1.2}$$

Note that any operator which has local symmetry with respect to H automatically also has global symmetry with respect to H but the converse implication does not necessarily hold. Indeed, generally the condition of local symmetry is much stronger than that of global symmetry.

The duality implies that within the symmetric and the antisymmetric subspaces of $(\mathbb{C}^d)^{\otimes n}$, the collective action of a gauge group G is dual to the collective action of G', its centralizer in $U(d)$. This in turn implies that if an operator is confined to the symmetric or antisymmetric subspace and has *global* symmetry with respect to the gauge group G, then it must also have *local* symmetry with respect to G. In other words, our generalization of Schur-Weyl duality allows in some circumstances for a global symmetry to be promoted to a local symmetry.

The main application we consider is the problem of how to best estimate parameters describing a quantum state given multiple copies of the state (this is considered in Sec. V). The parameters might consist of expectation values of some observables, or they might encode a decision about the state, such as whether a given expectation value is positive or not. In particular, we seek to determine under what circumstances it is sufficient to do independent measurements on each copy and in what circumstances more complicated measurements, for instance, using entanglement, are required. (As it turns out, there are many circumstances wherein entangled measurements do help.)

A very simple example of such a multi-copy estimation problem is the one considered by Hayashi *et al.* [3]. A pure state is chosen uniformly according to the Haar measure, and n copies of the state are prepared. The goal is to estimate the expectation value of an observable A for the state. Hayashi *et al.* have shown that for a squared-error figure of merit, the optimal estimation scheme is to simply measure the observable A separately on each system. Our generalization of Schur-Weyl duality can be used to provide a very elementary proof of this result. It can also be used to simplify the solution of estimation problems that are much more complicated, as we shall show.

The reason we can make use of our generalization of Schur-Weyl duality is that a multi-copy estimation problem can be shown to naturally have a global symmetry for a gauge group. Measurements with global symmetry relative to a gauge group are described by POVMs all the elements of which have this symmetry. In this case, the duality tells us that the global symmetry can be promoted to a local symmetry, so it suffices to consider measurements on the n copies that have *local symmetry* relative to the gauge group.

We can now explain how this promotion of global symmetries to local symmetries immediately leads to the solution of the multi-copy estimation problem considered by Hayashi *et al.*. The prior is uniform over pure states and the squared error figure of merit only relies on the observable A that one is trying to estimate. Consequently, the description of this problem has symmetry with respect to the group of all unitaries which commute with A and it follows that, on the multi-copy system, it suffices to perform measurements that have global symmetry with respect to this group. But this group is a gauge group, i.e. it is equal to the centralizer of its centralizer in the unitary group, and so by our result, one can promote this global symmetry to a local symmetry. Finally, noting that all measurement operators that are invariant under the local action of the gauge group must be in the algebra generated by the set

$$\{I^{\otimes k} \otimes A \otimes I^{\otimes(n-k-1)} : \; 0 \leq k \leq n-1\} \cup \{I^{\otimes n}\},$$

it follows that one can simply measure A on each copy individually and do classical processing on the outcomes to achieve the optimal estimation. We also immediately see that even if the figure of merit is not the squared error and the prior over pure states is not uniform, as long as these depend only on A, then an individual measurement on each copy continues to be optimal.

In more complicated examples, wherein there may be many observables to be estimated, a non-uniform prior over pure states and an arbitrary figure of merit, as long as the problem still has some gauge symmetry we can exploit our result to infer that the optimal measurement on the n copies should have local symmetry with respect to the gauge group.

In particular, if the commutant of the gauge group is a commutative algebra, then it suffices to implement independent measurements of the generator of this algebra on each system. This occurs if the problem is to estimate a set of commuting observables, and the figure of merit and the prior over pure states can be entirely described in terms of this set of observables (hence, no entanglement is needed). Furthermore, even if the commutant of the problem's gauge group is not a commutative alge-

bra, so that independent measurements are *not* sufficient, nonetheless local symmetry is still a stronger constraint than global symmetry and consequently our result can lead to a bound on how much interaction between the systems is required to achieve the optimal measurement.

We also demonstrate several other generalizations of the basic multi-copy estimation problem – to cases which include mixed states and to cases where the channel between the source and the estimator can be noisy – such that our results still have nontrivial consequences for the optimal measurement.

Finally, we demonstrate that our result has applications for estimation problems where the estimator gets only a single copy of the system of interest, and the distinction between global and local symmetries is relative to the partitioning of the system of interest into its components. In particular, we obtain strong constraints on the optimal measurement in the case of a system with *two* components because the permutation group on two systems has only irreducible representations over the symmetric and antisymmetric subspaces and our duality only permits an inference from global symmetry to local symmetry within the symmetric and antisymmetric subspaces. This is considered in Sec. VI.

Given that the class of estimation problems for which our results apply is very large, they represent a dramatic expansion, relative to previously known results, in the scope of problems for which we can easily determine the optimal measurement. Furthermore, in previous results where independent measurements on each copy were shown to be optimal, such as Ref. [3], the reasoning was rather *ad hoc*. It was not clear what feature of the estimation problem implied the sufficiency of such measurements. By contrast, our approach follows a clear methodology – we are determining the consequences of the gauge symmetries of the estimation problem. Our results establish a sufficient condition for the optimality of independent measurements, , i.e. the lack of any need for adaptive or entangled measurements. It is that the set of single-copy observables that are needed to define the estimation problem form a *commutative* set. In a slogan, *the commutativity of the observables defining the estimation problem imply the adequacy of independent measurements.*

II. PRELIMINARIES

A. Commutant and Centralizer

For a complex vector space $\mathcal{V}$, define $\mathrm{End}(\mathcal{V})$ to be the set of linear maps from $\mathcal{V}$ to itself (endomorphism). This set has a natural structure of algebra and is called the full matrix algebra over $\mathcal{V}$. Any matrix algebra defined on $\mathcal{V}$ is a subalgebra of $\mathrm{End}(\mathcal{V})$. Throughout this paper we only consider finite dimensional vector spaces.

For any vector space $\mathcal{V}$, and any set $\{A_i \in \mathrm{End}(\mathcal{V})\}$ we call the set of all operators in $\mathrm{End}(\mathcal{V})$ which commute with $\{A_i\}$ the *commutant* of $\{A_i\}$ and denote it by $\mathrm{Comm}\{A_i\}$. Note that for any arbitrary set $\{A_i \in \mathrm{End}(\mathcal{V})\}$, its commutant, i.e. $\mathrm{Comm}\{A_i\}$, is an algebra.

Let $\{A_i \in \mathrm{End}(\mathcal{V})\}$ be a set of Hermitian operators, i.e. $A_i = A_i^\dagger$. Then it holds that

$$\mathrm{Comm}\{\mathrm{Comm}\{A_i\}\} = \mathrm{Alg}\{A_i, I\} \tag{2.1}$$

where by $\mathrm{Alg}\{A_i, I\}$ we mean the complex matrix algebra generated by the set $\{A_i\}$ and I (identity operator on $\mathcal{V}$). Any such complex matrix algebra which includes identity operator and is closed under adjoint (†) is called a *finite dimensional von Neumann algebra*. Note that Eq.(2.1) means that for any finite dimensional von Neumann algebra $\mathcal{A}$

$$\mathrm{Comm}\{\mathrm{Comm}\{\mathcal{A}\}\} = \mathcal{A} \tag{2.2}$$

which is the defining property of these algebras. In this paper we only use this type of algebras and whenever we refer to an object as an algebra we mean a finite dimensional von Neumann algebra. Note that for any subgroup H of the unitary group the algebra spanned by H, $\mathrm{Alg}\{H\}$, is a von Neumann algebra.

A finite dimensional von Neumann algebra, as a finite dimensional matrix C*-algebra, has a unique decomposition up to unitary equivalence of the form

$$\mathcal{A} \cong \bigoplus_J (\mathcal{M}_{m_J} \otimes \mathbb{I}_{n_J}) \tag{2.3}$$

where $\mathcal{M}_{m_J}$ is the full matrix algebra $\mathrm{End}(\mathbb{C}^{m_J})$ and $\mathbb{I}_{n_J}$ is the identity on $\mathbb{C}^{n_J}$. A von Neumann algebra by definition includes identity. Therefore for these algebras $\sum_J m_J n_J$ is equal to the dimension of vector space.

For two algebras $\mathcal{A}_1 \subseteq \mathrm{End}(\mathcal{V}_1)$ and $\mathcal{A}_2 \subseteq \mathrm{End}(\mathcal{V}_2)$ it holds that

$$\mathrm{Comm}\{\mathcal{A}_1 \otimes \mathcal{A}_2\} = \mathrm{Comm}\{\mathcal{A}_1\} \otimes \mathrm{Comm}\{\mathcal{A}_2\} \tag{2.4}$$

this is called *commutation theorem for tensor products*.

In this paper we will use the notion of *centralizer* in a different way than *commutant*. By the *centralizer* of a subgroup H_0 in group H we mean the set of all elements of group H which commute with all elements of the subgroup H_0. We denote the centralizer of H_0 by H_0'. Note that the centralizer of any subgroup of a group is also a subgroup of that group.

Let H be a subgroup of $\mathrm{U}(d)$ and H' be its centralizer in this group. Then it holds that

$$\mathrm{Comm}\{H\} = \mathrm{Alg}\{H'\} \tag{2.5}$$

B. Dual reductive pairs and Schur-Weyl duality

Let H_1,H_2 be two groups of unitaries acting on the complex vector space $\mathcal{V}$ and assume that they commute with each other, that is, H_1 and H_2 are each within one

another's centralizer in the group of all unitaries on $\mathcal{V}$. Then, under the action of H_1 and H_2, the space $\mathcal{V}$ decomposes as follows

$$\mathcal{V} \cong \sum_{\mu,\nu} \mathcal{M}_\mu \otimes \mathcal{N}_\nu \otimes \mathbb{C}^{m_{\mu,\nu}} \tag{2.6}$$

where H_1 and H_2 act irreducibly on $\mathcal{M}_\mu$ and $\mathcal{N}_\mu$ respectively, where μ and ν label distinct irreducible representations (irreps) of H_1 and H_2 respectively and where $m_{\mu,\nu}$ is the multiplicity of irreps μ, ν. Then for some specific commuting groups the following equivalent properties hold [1, 2].

Proposition 1 *Let H_1,H_2 be two groups acting on $\mathcal{V}$. Then the followings are equivalent*

1. *The complex algebra spanned by H_1 is the commutant of the complex algebra spanned by H_2 in End($\mathcal{V}$) and vice versa.*

2. *In the decomposition 2.6 each $m_{\mu,\nu}$ is either 0 or 1 and at most one $m_{\mu,\nu}$ is nonzero for each μ and each ν.*

Any two groups with these properties are called a dual reductive pair *of subgroups of $GL(\mathcal{V})$ the general linear group on $\mathcal{V}$.*

Note that using the notation we have introduced before the first statement can be written as $\mathrm{Alg}\{H_1\} = \mathrm{Comm}\{H_2\}$ and by virtue of Eq.(2.2) this equation is equivalent to $\mathrm{Alg}\{H_2\} = \mathrm{Comm}\{H_1\}$.

Consider the following representation of the unitary group $\mathrm{U}(d)$ on $(\mathbb{C}^d)^{\otimes n}$.

$$\forall V \in \mathrm{U}(d): \quad \mathbf{Q}(V)|i_1\rangle \otimes \cdots \otimes |i_n\rangle = V|i_1\rangle \otimes \cdots \otimes V|i_n\rangle \tag{2.7}$$

For a subgroup H of $\mathrm{U}(d)$ we denote the group $\{\mathbf{Q}(V) : V \in H\}$ by $\mathbf{Q}(H)$ and we call it the *collective action* of H on $(\mathbb{C}^d)^{\otimes n}$. Consider also the *canonical representation* of the symmetric group of degree n, $\mathcal{S}_n$, on $(\mathbb{C}^d)^{\otimes n}$

$$\forall s \in \mathcal{S}_n: \quad \mathbf{P}(s)|i_1\rangle \otimes \cdots \otimes |i_n\rangle = |i_{s^{-1}(1)}\rangle \otimes \cdots \otimes |i_{s^{-1}(n)}\rangle \tag{2.8}$$

We denote the group $\{\mathbf{P}(s) : s \in \mathcal{S}_n\}$ by $\mathbf{P}(\mathcal{S}_n)$. Then Schur-Weyl duality states that

Theorem 2 ***(Schur-Weyl duality)*** *The following two algebras are commutants of one another in End($(\mathbb{C}^d)^{\otimes n}$)*

1. *Alg$\{\boldsymbol{Q}(U(d))\}$, the complex algebra spanned by $\boldsymbol{Q}(U(d))$.*

2. *Alg$\{\boldsymbol{P}(\mathcal{S}_n)\}$, the complex algebra spanned by $\boldsymbol{P}(\mathcal{S}_n)$.*

In other words, the subgroups $\boldsymbol{Q}(U(d))$ and $\boldsymbol{P}(\mathcal{S}_n)$ are dual reductive pairs in $GL((\mathbb{C}^d)^{\otimes n})$.

Using our notation Schur-Weyl duality can be expressed as $\mathrm{Comm}\{\mathbf{Q}(\mathrm{U}(d))\} = \mathrm{Alg}\{\mathbf{P}(\mathcal{S}_n)\}$ or equivalently as $\mathrm{Alg}\{\mathbf{Q}(\mathrm{U}(d))\} = \mathrm{Comm}\{\mathbf{P}(\mathcal{S}_n)\}$.

This theorem together with the proposition 1 implies that there is a one-to-one correspondence between the irreps of the group $\mathrm{U}(d)$ which show up in representation $\mathbf{Q}(\mathrm{U}(d))$ and the irreps of the group $\mathcal{S}_n$ which show up in representation $\mathbf{P}(\mathcal{S}_n)$. Furthermore, the theorem implies that the action of $\mathbf{Q}(\mathrm{U}(d)) \times \mathbf{P}(\mathcal{S}_n)$ is multiplicity-free on $(\mathbb{C}^d)^{\otimes n}$.

In the following section, we present a generalization of Schur-Weyl duality for the case of *gauge* subgroups of U(d).

III. A GENERALIZATION OF SCHUR-WEYL DUALITY

A. Gauge groups and their characterizations

For any subgroup G of $\mathrm{U}(d)$ let G' denotes the centralizer of G in $\mathrm{U}(d)$ i.e. the set of all elements of $\mathrm{U}(d)$ which commutes with all elements of G. Also denote the centralizer of the centralizer of G by $G'' \equiv (G')'$. Then in general $G \subseteq G''$. We call a unitary group G a *gauge group* if $G = G''$. The fact that in any arbitrary group and for any arbitrary subgroup H, $H \subseteq H''$ implies that $((H')')' = H'$. So for arbitrary subgroup H of $\mathrm{U}(d)$, its centralizer H' is a gauge group.

Equivalently, one can think of a gauge group as the set of all unitaries in $\mathrm{End}(\mathbb{C}^d)$ which commute with a von Neumann algebra $\mathcal{A} \subseteq \mathrm{End}(\mathbb{C}^d)$. This is true because for any subgroup G of $\mathrm{U}(d)$, G'' is equal to all the unitries which commute with G' or equivalently all the unitaries which commute with $\mathrm{Alg}\{G'\}$ (which is a von Neumann algebra). So if $G = G''$ then G is equal to the set of all unitaries which commute with an algebra, namely $\mathrm{Alg}\{G'\}$. On the other hand, if G is equal to the set of all unitaries in $\mathrm{End}(\mathbb{C}^d)$ which commute with an algebra $\mathcal{A} \subseteq \mathrm{End}(\mathbb{C}^d)$ then G' is equal to all the unitaries in the algebra $\mathcal{A}$ and so is a basis for this algebra. Since G'' is equal to the set of all the unitaries which commute with G', and G' is a basis for $\mathcal{A}$, then G'' is equal to the set of all unitaries which commute with the algebra $\mathcal{A}$ and so is equal to G. Therefore, these two definitions of gauge group are equivalent.

This discussion implies that one way to specify a gauge group is to specify the von Neumann algebra of operators which commute with the gauge group, for instance by specifying the generators of that algebra. We call the gauge group formed by all unitaries which commute with a von Neumann algebra $\mathcal{A}$ *the gauge group of* $\mathcal{A}$ and denote it by $G_{\mathcal{A}}$. Note that if $G_{\mathcal{A}}$ is the gauge group of $\mathcal{A}$ then it holds that

$$\mathrm{Comm}\{G_{\mathcal{A}}\} = \mathrm{Alg}\{G'_{\mathcal{A}}\} = \mathcal{A}. \tag{3.1}$$

Using this together with commutation theorem for tensor product Eq.(2.4) and Eq.(2.5) we find

$$\mathrm{Comm}\{G_{\mathcal{A}}^{\times n}\} = \mathrm{Alg}\{(G'_{\mathcal{A}})^{\times n}\} = \mathcal{A}^{\otimes n} \tag{3.2}$$

Also note that Eq.(3.1) implies that any von Neumann algebra can be uniquely specified by its gauge group.

Now based on this observation that any gauge group can be thought as the set of unitaries commuting with a von Neumann algebra characterizing the set of all gauge groups is equivalent to characterizing all von Neumann algebras which is done by Eq.(2.3). This decomposition implies that $G_{\mathcal{A}}$ the gauge group of $\mathcal{A}$, has a unique decomposition up to unitary equivalence of the form

$$G_{\mathcal{A}} \cong \bigoplus_J (\mathbb{I}_{m_J} \otimes \mathrm{U}(n_J)) \tag{3.3}$$

where $\mathbb{I}_{m_J}$ is identity on $\mathbb{C}^{m_J}$ and $\sum_J n_J m_J = d$. In other words, for any set of integers $1 \leq n_1 \leq \cdots \leq n_k \leq d$ there is a gauge group acting on $\mathbb{C}^d$ which is isomorphic to $\mathrm{U}(n_1) \times \cdots \times \mathrm{U}(n_k)$ iff there is a set of positive integers $1 \leq m_1, \cdots, m_k \leq d$ such that $\sum_i n_i m_i = d$. In particular, for any vector space $\mathbb{C}^d$, there are gauge groups isomorphic to $\mathrm{U}(1)^{\times d}$ and $\mathrm{U}(d)$. These gauge groups can be respectively thought as the gauge group of the algebra of all diagonal matrices in some orthonormal basis and the algebra generated by identity matrix.

For instance, in the case of $d = 2$ the set of all gauge groups can be classified in the following three types: i) $n_1 = 0, n_2 = 1$ which corresponds to the group $\{e^{i\theta} I : \theta \in (0, 2\pi]\}$ where I is the identity operator , ii) $n_1 = 0, n_2 = 2$ which corresponds to the group $\mathrm{U}(2)$ iii) $n_1 = 1, n_2 = 1$ which corresponds to the group

$$\{e^{i\theta_0}|0\rangle\langle 0| + e^{i\theta_1}|1\rangle\langle 1| : \theta_0, \theta_1 \in (0, 2\pi]\}$$

for any arbitrary orthonormal basis $\{|0\rangle, |1\rangle\}$.

Note that this characterization implies that any nontrivial gauge group is a unimodular Lie group, i.e. its left invariant measure is equal to the right invariant measure (up to a constant).

Throughout this paper we will extensively use the uniform twirling over subgroups of unitary group with respect to unique (normalized) Haar measure of H denoted by

$$\mathcal{T}_H(\cdot) \equiv \int_H d\mu(V)\ V(\cdot)V^\dagger \tag{3.4}$$

where $d\mu$ is the normalized Haar measure of H. Since $d\mu$ is the uniform measure any operator in the image of $\mathcal{T}_H$ commutes with H. Therefore if $G_{\mathcal{A}}$ is the gauge group of a von Neumann algebra $\mathcal{A}$ then the image of $\mathcal{T}_{G_{\mathcal{A}}}$ is inside the algebra $\mathcal{A}$.

Finally, it is worth noting that if G is a gauge group then the two groups G and G' are dual reductive pairs. However, the inverse is not true, i.e. if two groups are dual reductive pairs, they are not necessarily each other's centralizers in the group of all unitaries. For example, according to the Schur-Weyl duality, the canonical representation of the permutation group on $(\mathbb{C}^d)^{\otimes n}$, i.e. $\mathbf{P}(\mathcal{S}_n)$, and the collective action of $\mathrm{U}(d)$, i.e. $\mathbf{Q}(\mathrm{U}(d))$, are dual reductive pairs but they are surely not equal to one another's centralizer in the group of all unitares acting on $(\mathbb{C}^d)^{\otimes n}$.

B. From gauge groups to dual reductive pair on product spaces

For a subgroup H of $\mathrm{U}(d)$ we denote $H^{\times n}$ to be the group $H^{\times n} \equiv \{U_1 \otimes \cdots \otimes U_n : U_i \in H\}$. Also, let $\langle H^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$ denote the group acting on $(\mathbb{C}^d)^{\otimes n}$ which is generated by the two groups $H^{\times n}$ and $\mathbf{P}(\mathcal{S}_n) = \{\mathbf{P}(s) : s \in \mathcal{S}_n\}$. Note that every element of $\langle H^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$ can be written in the canonical form of $W\mathbf{P}(s)$ for a unique $W \in H^{\times n}$ and a unique $s \in \mathcal{S}_n$. This implies a homomorphism from $\langle H^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$ to $\mathbf{P}(\mathcal{S}_n)$ with the kernel $H^{\times n}$, and therefore $\langle H^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle = H^{\times n} \rtimes \mathbf{P}(\mathcal{S}_n)$.

Then one can prove the following generalization of Schur-Weyl duality

Theorem 3 ***(Generalization of Schur-Weyl duality)*** *Suppose G and G' are one another's centralizers in the group of unitaries U(d). Then the following two algebras are commutants of one another in End$((\mathbb{C}^d)^{\otimes n})$*

1. *Alg$\{\boldsymbol{Q}(G)\}$, the complex algebra spanned by $\boldsymbol{Q}(G)$.*
2. *Alg$\{(G')^{\times n}, \boldsymbol{P}(\mathcal{S}_n)\}$, the complex algebra spanned by $\langle (G')^{\times n}, \boldsymbol{P}(\mathcal{S}_n)\rangle$.*

In other words, the subgroups $\boldsymbol{Q}(G)$ and $\langle (G')^{\times n}, \boldsymbol{P}(\mathcal{S}_n)\rangle$ are dual reductive pairs in GL$((\mathbb{C}^d)^{\otimes n})$.

Using Eq.(3.2) we can rephrase the theorem as

Corollary 4 *Let $G_{\mathcal{A}}$ be the gauge group of the von Neumann algebra $\mathcal{A} \subseteq$ End$(\mathbb{C}^d)$. Then*

$$\mathit{Comm}\{\boldsymbol{Q}(G_{\mathcal{A}})\} = \mathit{Alg}\{\mathcal{A}^{\otimes n}, \boldsymbol{P}(\mathcal{S}_n)\}\,. \tag{3.5}$$

This form of theorem is particularly useful and has a straightforward physical interpretation which we will point out in the next section and discuss in more details in [7].

Theorem 3 together with the proposition 1 implies that there is a one-to-one correspondence between the irreps of the group G which show up in representation $\mathbf{Q}(G)$ on $(\mathbb{C}^d)^{\otimes n}$ and the irreps of the group $\langle (G')^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$ which show up in this space. Furthermore, the theorem implies that the representation of $\mathbf{Q}(G) \times \langle (G')^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$ is multiplicity-free on $(\mathbb{C}^d)^{\otimes n}$. Note that in the specific case of $G = \mathrm{U}(d)$ (where G' is the trivial group) this dual reductive pair reduces to the well-known Schur-Weyl duality.

Also note that the fact that each of the algebras in this theorem is in the commutant of the other algebra is

trivial. In other words, for any subgroup $H \subseteq \mathrm{U}(d)$ it holds that

$$\mathrm{Alg}\{\mathbf{Q}(H)\} \subseteq \mathrm{Comm}\{(H')^{\times n}, \mathbf{P}(\mathcal{S}_n)\}$$

The non-trivial content of the theorem is that for gauge groups these two algebras are equal. For H a subgroup of $\mathrm{U}(d)$ that is not equal to its bicommutant in $\mathrm{U}(d)$, and so is not a gauge group, the above two algebras are not necessarily equal. We provide a simple example illustrating this fact in Appendix B.

To prove theorem 3 we use the following property of gauge groups which is proven in Appendix A..

Lemma 5 *For a gauge group G, the complex algebra spanned by $\boldsymbol{Q}(G)$ is equal to the permutationally invariant subalgebra of the complex algebra spanned by $G^{\times n}$.*

The result can be summarized as

$$\begin{aligned} G'' = G \;\; \rightarrow \;\; \mathrm{Alg}\{\mathbf{Q}(G)\} &= \mathrm{Alg}\{G^{\times n}\} \;\cap\; \mathrm{Comm}\{\mathbf{P}(\mathcal{S}_n)\} \\ &= \mathrm{Alg}\{G\}^{\otimes n} \;\cap\; \mathrm{Comm}\{\mathbf{P}(\mathcal{S}_n)\}. \end{aligned}$$

Using this lemma the proof of theorem 3 is then straightforward and proceeds as follows.

Proof. (**Theorem** 3) Since both algebras are von Neumann algebra, we only need to show that one is the commutant of the other, the other direction follows from Eq.(2.2). So to prove the theorem it is sufficient to show that $\mathrm{Comm}\{G'^{\times n}, \mathbf{P}(\mathcal{S}_n)\} = \mathrm{Alg}(\mathbf{Q}(G))$. To show this, we note that

$$\mathrm{Comm}\{G'^{\times n}, \mathbf{P}(\mathcal{S}_n)\} = \mathrm{Comm}\{G'^{\times n}\} \cap \mathrm{Comm}\{\mathbf{P}(\mathcal{S}_n)\}.$$

Then since $\mathrm{Comm}\{G'^{\times n}\} = \mathrm{Alg}\{G^{\times n}\}$ we conclude that

$$\mathrm{Comm}\{G'^{\times n}, \mathbf{P}(\mathcal{S}_n)\} = \mathrm{Alg}\{G^{\times n}\} \;\cap \mathrm{Comm}\{\mathbf{P}(\mathcal{S}_n)\}.$$

This together with lemma 5 completes the proof of theorem. ■

Finally, it is worth mentioning the following corollary of lemma 5 which applies to arbitrary subgroup of $\mathrm{U}(d)$

Corollary 6 *For any unitary subgroup $H \subseteq U(d)$, permutationally invariant subalgebra of Comm$\{H^{\times n}\}$ is equal to Alg$\{\boldsymbol{Q}(H')\}$.*

Proof. First note that Eq.(2.5) together with commutation theorem for tensor products, i.e. Eq.(2.4), implies

$$\mathrm{Comm}\{H^{\times n}\} = \mathrm{Alg}\{(H')^{\times n}\}$$

Then, from section III A we know that the centralizer of H an arbitrary subgroup of $\mathrm{U}(d)$ is a gauge group and so one can apply lemma 5 for gauge group H' which implies that the permutationally invariant subalgebra of $\mathrm{Alg}\{(H')^{\times n}\}$ is equal to $\mathrm{Alg}\{\mathbf{Q}(H')\}$. ■

C. An intuitive account

Our generalization of Schur-Weyl duality appears very intuitive if one considers a particular problem concerning two independent observers using different conventions to describe quantum systems.

Suppose that Alice and Bob each use their own personal convention to associate observables with operators in the Hilbert space of a system, and assume that each observer is not aware of the other's convention. All they know is that for a particular set of operators $\{A_i\}$, the observable which is described by operator A_i relative to Alice's convention is also described by operator A_i relative to Bob's convention. Clearly, Alice and Bob will also agree on any observable which is an algebraic function of the $\{A_i\}$ and the identity operator I, so the full set of observables on which they agree are those in the algebra $\mathcal{A} \equiv \mathrm{Alg}\{A_i, I\}$. The question is: what sorts of states and observables can they agree upon for the composite of n systems, assuming Alice and Bob use the same convention for each system and agree on how to label the systems?

It is obvious that Alice and Bob agree on the description of all observables which are in the algebra generated by (i) the n-fold tensor product of the algebra $\mathcal{A}$ and (ii) the canonical representation of the permutation group. Furthermore, it is intuitively clear that there are no other observables in addition to these that they can agree upon.

Now note that the group $G_{\mathcal{A}}$ of unitaries that commute with $\mathcal{A}$ can be interpreted as the possible ways in which Alice and Bob's conventions may be related to one another. Because Alice and Bob use the *same* convention for each of the n systems, the operators that they can agree on for the composite are those that are invariant under the *collective* action of $G_{\mathcal{A}}$, i.e. $\mathbf{Q}(G_{\mathcal{A}})$. What is intuitively clear, therefore, is that the operators that are in the commutant of the collective action of $G_{\mathcal{A}}$ are those in the algebra spanned by the n-fold product of $\mathcal{A}$, $\mathcal{A}^{\otimes n}$, and the canonical representation of the permutation group, $\mathbf{P}(\mathcal{S}_n)$. But this is precisely the content of our generalization of Schur-Weyl duality, in the form presented in corollary 6. (Indeed, it was in attempting to make this intuition rigorous that we were led to prove the duality.) We discuss more on this physical interpretation of our generalization of Schur-Weyl duality in [7].

This discussion also reveals the motivation for calling the group $G_{\mathcal{A}}$ a *gauge group*. It is because such a group describes the possible transformations that leave the physically relevant set of observables invariant (in this case, the single-system observables that Alice and Bob agree upon), and such transformations are typically called gauge transformations by physicists.

D. Duality within the symmetric and antisymmetric subspaces

In the special case where the support of operators are restricted to the symmetric or anti-symmetric subspace, theorem 3 has an interesting corollary. Let $\Pi_\pm$ be the projector to $\left[(\mathbb{C}^d)^{\otimes n}\right]_\pm$, the symmetric (respectively antisymmetric) subspace of $(\mathbb{C}^d)^{\otimes n}$. Then we can prove that

Theorem 7 *Suppose G and G' are one another's centralizers in the group of unitaries U(d). Then the following two algebras are the commutants of one another in End$(\left[(\mathbb{C}^d)^{\otimes n}\right]_\pm)$*

1. *Alg$\{\Pi_\pm \boldsymbol{Q}(G)\Pi_\pm\}$, the complex algebra spanned by $\Pi_\pm \boldsymbol{Q}(G)\Pi_\pm$.*

2. *Alg$\{\Pi_\pm \boldsymbol{Q}(G')\Pi_\pm\}$, the complex algebra spanned by $\Pi_\pm \boldsymbol{Q}(G')\Pi_\pm$.*

In other words, $\Pi_\pm \boldsymbol{Q}(G)\Pi_\pm$ and $\Pi_\pm \boldsymbol{Q}(G')\Pi_\pm$ are dual reductive pairs in GL$(\left[(\mathbb{C}^d)^{\otimes n}\right]_\pm)$.

Again, the fact that each of these algebras is in the commutant of the other is trivial. The non-trivial fact is that each is *equal* to the commutant of the other. We can summarize the theorem by

$$G'' = G \quad \Longrightarrow \quad \mathrm{Comm}\{\Pi_\pm \mathbf{Q}(G)\Pi_\pm\} = \mathrm{Alg}\{\Pi_\pm \mathbf{Q}(G')\Pi_\pm\}. \tag{3.6}$$

where here by $\mathrm{Comm}\{\Pi_\pm \mathbf{Q}(G)\Pi_\pm\}$ we mean the set of all operators in $\mathrm{End}(\left[(\mathbb{C}^d)^{\otimes n}\right]_\pm)$ which commute with $\Pi_\pm \mathbf{Q}(G)\Pi_\pm$.

Proof. (**Theorem** 7) Again since both algebras are von Neumann algebra, we only need to show that $\mathrm{Comm}\{\Pi_\pm \mathbf{Q}(G')\Pi_\pm\} = \mathrm{Alg}\{\Pi_\pm \mathbf{Q}(G)\Pi_\pm\}$. Let M be an arbitrary operator in $\mathrm{End}((\mathbb{C}^d)^{\otimes n})$ such that $\Pi_\pm M \Pi_\pm$ commutes with $\Pi_\pm \mathbf{Q}(G')\Pi_\pm$. Then $\Pi_\pm M \Pi_\pm$ clearly commutes with $\mathbf{Q}(G')$ and therefore theorem 3 implies that it is in the span of $\langle G^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$. Now recall that, every arbitrary element of $\langle G^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$ can be written in the canonical form of $W\mathbf{P}(s)$ for a unique $W \in G^{\times n}$ and a unique $s \in \mathcal{S}_n$. So

$$\Pi_\pm M \Pi_\pm = \sum_{W\in G^{\times n},\ s\in\mathcal{S}_n} c_{W,s}\ \ W\mathbf{P}(s) \tag{3.7}$$

for some complex coefficients $c_{W,s}$. Then

$$\Pi_\pm M \Pi_\pm = \Pi_\pm \left[\sum_{W\in G^{\times n},\ s\in\mathcal{S}_n} (-1)^{p_\pm(s)}\ c_{W,s}\ \ W\right] \Pi_\pm \tag{3.8}$$

where $\mathbf{P}(s)\Pi_\pm = (-1)^{p_\pm(s)}\Pi_\pm$ for arbitrary $s \in \mathcal{S}_n$, $(-1)^{p_+(s)} = 1$ for all $s \in \mathcal{S}_n$ and $(-1)^{p_-(s)} = \pm 1$ dependent on whether s is an odd or even permutation. Therefore there exist an operator $\bar{M}$ in the span of $G^{\times n}$ such that $\Pi_\pm \bar{M} \Pi_\pm = \Pi_\pm M \Pi_\pm$. Then

$$\Pi_\pm \left[\sum_{s\in\mathcal{S}_n} \mathbf{P}(s)\bar{M}\mathbf{P}^\dagger(s)\right] \Pi_\pm = \Pi_\pm \bar{M} \Pi_\pm = \Pi_\pm M \Pi_\pm \tag{3.9}$$

where we have used the fact that $\Pi_\pm \mathbf{P}(s) = \mathbf{P}^\dagger(s)\Pi_\pm = (-1)^{p_\pm(s)}\Pi_\pm$ and two negative signs cancel each other. Since $\bar{M}$ is in the span of $G^{\times n}$ then $\tilde{M} \equiv \sum_{s\in\mathcal{S}_n} \mathbf{P}(s)\bar{M}\mathbf{P}^\dagger(s)$ is in the permutationally invariant subalgebra of the span $G^{\times n}$. Now since G is gauge group using lemma 5 we can conclude that $\tilde{M} \in \mathrm{Alg}\{\mathbf{Q}(G)\}$. So for any arbitrary $M \in \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ if $\Pi_\pm M \Pi_\pm$ commutes with $\Pi_\pm \mathbf{Q}(G')\Pi_\pm$ then there exists an operator $\tilde{M}$ in $\mathrm{Alg}\{\mathbf{Q}(G)\}$ such that $\Pi_\pm \tilde{M} \Pi_\pm = \Pi_\pm M \Pi_\pm$. This completes the proof of theorem. ■

Again, using the proposition 1 one can see that theorem 7 implies: (i) a one-to-one correspondence between the irreps of G which show up in the representation $\mathbf{Q}(G)$ in the symmetric (antisymmetric) subspace and the irreps of G' which show up in the representation $\mathbf{Q}(G')$ in the symmetric (antisymmetric) subspace, and (ii) that in these subspaces $\mathbf{Q}(G) \times \mathbf{Q}(G')$ is multiplicity-free. The special case of this result is known in the representation theory for the case of symmetric subspace of $(\mathbb{C}^{d_1 d_2})^{\otimes n}$ and the collective representation of $G = \mathrm{U}(d_1) \times e$ and $G' = e \times \mathrm{U}(d_2)$ as two subgroups of $\mathrm{U}(d_1 d_2)$

Applying theorem 7 for $G_\mathcal{A}$ the gauge group of a von Neumann algebra $\mathcal{A}$ one can show that for any given operator $\Pi_\pm M \Pi_\pm$ which commutes with $\mathbf{Q}(G_\mathcal{A})$ there is an operator $\tilde{M}_\pm$ in the permutationally invariant subalgebra of $\mathcal{A}^{\otimes n}$ such that

$$\Pi_\pm \tilde{M}_\pm \Pi_\pm = \Pi_\pm M \Pi_\pm.$$

However, this argument is not constructive and for a given M it is not clear how we can find such an operator $\tilde{M}_\pm$ with this property. In the following theorem, we introduce a completely positive unital quantum operation which does this transformation.

Theorem 8 *Let $G_\mathcal{A} \subseteq U(d)$ be the gauge group of a von Neumann algebra $\mathcal{A} \subseteq$ End$(\mathbb{C}^d)$. Then there exists a superoperator $\mathcal{L}_\pm$ from End$\left((\mathbb{C}^d)^{\otimes n}\right)$ to itself such that*

1. *$\mathcal{L}_\pm$ is unital and completely positive,*

2. *The image of $\mathcal{L}_\pm$ is in permutationally invariant subalgebra of $\mathcal{A}^{\otimes n}$ and*

3. *if $\Pi_\pm M \Pi_\pm$ commutes with $\boldsymbol{Q}(G_\mathcal{A})$ then*

$$\Pi_\pm \mathcal{L}_\pm(M) \Pi_\pm = \Pi_\pm M \Pi_\pm$$

An instance of such a superoperator is given by

$$\mathcal{L}_\pm(\cdot) \equiv \Phi_\pm(\cdot) + \frac{\mathrm{tr}(\cdot)}{d^n}\left[I^{\otimes n} - \Phi_\pm(I^{\otimes n})\right], \tag{3.10}$$

with

$$\Phi_{\pm}(\cdot) \equiv \bigoplus_{\mu} p_{\mu,\pm}^{-1}\ P_\mu[\mathcal{T}_{G'_{\mathcal{A}}}^{\otimes n}(\Pi_{\pm}(\cdot)\Pi_{\pm})]P_\mu \qquad (3.11)$$

where μ labels all the irreps of $G'_{\mathcal{A}}$ which show up in the representation $\boldsymbol{Q}(G'_{\mathcal{A}})$, P_μ is the projector to the subspace of $(\mathbb{C}^d)^{\otimes n}$ associated to irrep μ, $p_{\mu,\pm} \equiv tr\left(P_\mu \mathcal{T}_{G'_{\mathcal{A}}}^{\otimes n}(\Pi_{\pm})\right)$ and the summation in Eq. (3.11) is over all the irreps μ for which p_μ is nonzero.

This is proven in appendix A. This theorem will be particularly useful in the rest of this paper.

Finally, it is worth to emphasize on the importance of the restriction to symmetric and anti-symmetric subspaces in the results of this section. As we have argued before, the fact that G is a gauge group plays a crucial role in theorem 3 (and therefore theorems 7 and 8). Similarly one can also show that the restriction to the symmetric or anti-symmetric subspace is also important for theorem 7 and 8 to hold. Let λ labels different irreps of the permutation group and Π_λ be the projector to the subspace $\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda$ of $(\mathbb{C}^d)^{\otimes n}$ in which the representation $\mathbf{P}(\mathcal{S}_n)$ acts like the irrep λ of $\mathcal{S}_n$. The goal is to see whether in theorem 7 one can substitute the projection to the symmetric (anti-symmetric) subspace by the projection to an arbitrary irrep λ of $\mathcal{S}_n$. However, for subspaces other than symmetric and anti-symmetric subspaces $\Pi_\lambda M \Pi_\lambda$ for arbitrary $M \in \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ is not necessarily permutationally invariant while the span of $\mathbf{Q}(G')$ is permutationally invariant. So to generalize theorem 7 to other representations of the permutation group one should make an extra assumption to guarantee that $\Pi_\lambda M \Pi_\lambda$ is also permutationally invariant. Then one may expect the following to be true: a permutationally invariant operator $\Pi_\lambda M \Pi_\lambda$ which commutes with $\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda$ is in the span of $\Pi_\lambda \mathbf{Q}(G')\Pi_\lambda$. However, as the example in appendix C explicitly shows, this conjecture is wrong, i.e.

$$\mathrm{Alg}\{\Pi_\lambda \mathbf{Q}(G')\Pi_\lambda\} \neq \mathrm{Comm}\{\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda\} \cap \mathrm{Comm}\{\Pi_\lambda \mathbf{P}(\mathcal{S}_n)\Pi_\lambda\}$$

where the commutant is defined in $\mathrm{End}(\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda)$. (Note that still the algebra in the left-hand side is a subalgebra of the algebra in the right-hand side.) In other words, the restriction to the symmetric and anti-symmetric subspace plays an important role in theorem 7 and 8.

IV. GENERAL APPLICATIONS IN QUANTUM INFORMATION

Schur-Weyl duality has many applications in quantum information theory and so we expect that this generalization will as well. Here we present two specific important examples of these applications. The first example is about finding noiseless subystems for collective noise associated with a gauge group, and the second is about how, for n copies of a system in a pure state confined to the symmetric or antisymmetric subspace, a measurement with global symmetry relative to a gauge group can be simulated by one that has local symmetry for that group. This second result is the seed of the next section, where we will consider the consequences for multi-copy estimation problems in more depth.

A. Characterizing the multi-partite operators that are globally symmetric

Many applications of Schur-Weyl duality in quantum information theory are based on the fact that it provides a simple characterization of all operators in $\mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ which commute with $\mathbf{Q}(\mathrm{U}(d))$ or conversely all operators in $\mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ which commute with $\mathbf{P}(\mathcal{S}_n)$.

Theorem 3 and its corollary 4 immediately yield a characterization of operators with global symmetry under a gauge group G – they lie in the span of the local action of G' and the action of the permutation group. Similarly, corollary 7 yields a characterization of operators confined to the symmetric and antisymmetric subspaces that have global symmetry under G. These are simply the operators in the span of the collective action of G'.

A straightforward application of this characterization is to find noiseless subsystems. In the following we present a simple example of this.

1. Example: Finding noiseless subsystems

We begin by reviewing the standard story about noiseless subsystems. Suppose one is going to send quantum information through a noisy qubit channel, where the noise is described by a unitary that is sampled at random, but wherein the same unitary acts on each qubit. For example, the qubits could be spin 1/2 particles with a nonzero mangetic moment and the noise could be due to a random magnetic field. As another example, the qubits could be realized as the polarization of photons sent through a fiber-optic cable and the noise could be due to random strains in the cable that induce changes in the polarization. In many cases, it is a good approximation to assume that the noise varies slowly compared to the interval between the qubits as they pass down the channel (or that it varies little on the distance scale between the qubits in the case of a quantum memory), in which case one can assume that the same random unitary is applied to all n qubits. Then it turns out that, due to the symmetry of the noise, it is possible to encode classical and quantum information in the n qubit system in such a way that it remains unaffected by the noise[8–10]. To see this, note that under these assumptions, the noise is described by the group $\mathbf{Q}(\mathrm{U}(2))$. Any state in the commutant of $\mathbf{Q}(\mathrm{U}(2))$ is invariant under

the noise. Furthermore, any state in the span of $\mathbf{P}(\mathcal{S}_n)$ has this property as well. Now using Schur-Weyl duality one can conclude that the span of $\mathbf{P}(\mathcal{S}_n)$ is equal to the commutant of $\mathbf{Q}(\mathrm{U}(2))$ and therefore *every* state which is unaffected by this type of noise is in the span of $\mathbf{P}(\mathcal{S}_n)$.

In a more general model, the system sent through the channel may have other degrees of freedom which can potentially be used to send quantum information. In other words, the Hilbert space describing each particle sent through the channel is not $\mathbb{C}^2$ but it is $\mathbb{C}^2 \otimes \mathcal{H}$ where the finite dimensional Hilbert space $\mathcal{H}$ describes another degree of freedom which is invariant under the noise in the channel. For example, in the case of photons one can use time-bin encoding in addition to the polarization encoding to encode an extra qubit in each photon. But the time-bin qubit does not suffer from depolarization or polarization mode-dispersion. In other words, this degree of freedom is invariant under polarization noise.

So we assume the noise in the channel is described by a random unitary in the form of $V \otimes \mathbb{I}_{\mathcal{H}}$ where $V \in \mathrm{U}(2)$ and it acts on $\mathbb{C}^2$ and $\mathbb{I}_{\mathcal{H}}$ is the identity operator acting on $\mathcal{H}$. In the case of a single system sent through the channel ($n = 1$), it is clear that any information encoded in the subsystem $\mathcal{H}$ is preserved under this type of noise. Consider the case of many systems sent through the channel ($n > 1$). The question is what are the set of all states of the n systems which are invariant under this type of noise. In other words, what is the set of all states which commute with $\mathbf{Q}(\mathrm{U}(2) \otimes \mathbb{I}_{\mathcal{H}})$? Clearly, in this case, the usual form of Schr-Weyl duality does not apply. But one can use the generalization of Schur-duality we presented in the previous section to find these density operators.

To see this, first note that the group of unitaries $G \equiv \{V \otimes \mathbb{I}_{\mathcal{H}} : V \in \mathrm{U}(2)\}$ is the gauge group of the algebra $\mathbb{I}_2 \otimes \mathrm{End}(\mathcal{H})$ where $\mathbb{I}_2$ is the identity operator on $\mathbb{C}^2$. Then corollary 4 (which is indeed another version of theorem 3) gives the characterization of all operators which commutes with $\mathbf{Q}(G)$: That is basically the set of all operators in $\mathrm{Alg}\{\mathbf{P}(\mathcal{S}_n), (\mathbb{I}_2 \otimes \mathrm{End}(\mathcal{H}))^{\otimes n}\}$. So the set of all density operators in this algebra is exactly the set of all states which remains unaffected under this type of noise. This means that to protect information one needs to encode it in either the invariant degree of freedom of each subsystem ($\mathcal{H}$) or in the permutational degree of freedom. Again note that even without using our results it is straightforward to see that all of these states remains unchanged under this noise. The non-trivial consequence of corollary is that this algebra includes all such density operators.

Note that if the group $H \subseteq \mathrm{U}(d)$ describing the noise is not a gauge group then the $\mathrm{Comm}\{\mathbf{Q}(H)\}$ can be larger than $\mathrm{Alg}\{(H')^{\otimes n}, \mathbf{P}(\mathcal{S}_n)\}$ as it is shown by a simple example in Appendix B (where the group H is the $j = 1$ representation of SU(2) in $\mathbb{C}^3$). This means that, unlike the case of noise described by a gauge group, one can encode quantum information in a space which is larger than the permutational degree of freedom of the systems together with the invariant degrees of freedom of each system.

B. Promoting global symmetries to local symmetries

Another important application of this new duality is that in particular cases one can *promote a global symmetry to local symmetry* as we will describe in this section.

Recall the definition of local and global symmetries for an arbitrary operator $M \in \mathrm{End}(\mathbb{C}^d)^{\otimes n}$. M has a global symmetry with respect to the symmetry group $H \subseteq \mathrm{U}(d)$ if M is invariant under the collective action of H, as specified in Eq. (1.1), i.e. if $M \in \mathrm{Comm}\{\mathbf{Q}(H)\}$. Similarly we say that M has local symmetry with respect to the symmetry group H if it is invariant under the local action of H, as specified in Eq. (1.2), i.e. if $M \in \mathrm{Comm}(H^{\times n})$.

As noted in the introduction, the condition of local symmetry is generally much stronger than global symmetry. For example, if H is the group of rotations then global symmetry of a Hamiltonian with respect to H implies that the vector of the total angular momentum of n systems is a constant of the motion in the dynamics generated by this Hamiltonian. However, in this case the angular momenta of the subsystems are not necessarily conserved and the n subsystems can exchange angular momentum with one another. On the other hand, having a Hamiltonian with local symmetry with respect to the group of rotation implies the existence of non-trivial constants of motion defined on each of the n subsystems. So in this case we will have n conserved vectors of angular momentums and under this type of Hamiltonian, subsystems cannot exchange angular momentum.

Now consider the case where the symmetry under consideration is described by a gauge group $G_{\mathcal{A}}$ of a von Neumann algebra $\mathcal{A} \subseteq \mathrm{End}(\mathbb{C}^d)$. Note that if $M \in \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ has global symmetry with respect to $G_{\mathcal{A}}$ then $\Pi_\pm M \Pi_\pm$ will also have global symmetry with respect to $G_{\mathcal{A}}$. Then according to theorem 8 for any operator M with global symmetry with respect to $G_{\mathcal{A}}$ there is an operator $\tilde{M}_\pm$ which has local symmetry with respect to $G_{\mathcal{A}}$ and is equal to M within the symmetric (anti-symmetric) subspace,

$$\Pi_\pm \tilde{M}_\pm \Pi_\pm = \Pi_\pm M \Pi_\pm$$

One can choose $\tilde{M}_\pm = \mathcal{L}_\pm(M)$ where $\mathcal{L}_\pm$ is the completely positive, unital superoperator defined in theorem 8. So using the terminology of local and global symmetry we can interpret theorem 8 as *promoting global symmetry to local symmetry.*

In the following we explore the important consequence of promoting global symmetry to local symmetry for the case of measurements.

1. From global to local symmetry for Measurements

The most general type of measurement that can be performed on a quantum system can be described by a POVM (positive operator-valued measure) (See e.g. [4, 5]). Consider a POVM $M : \sigma(\Omega) \to \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$. Here, Ω denotes the space of outcomes of the measurement. This is a measure space equipped with a σ-algebra of subsets, denoted by $\sigma(\Omega)$. The elements of the σ-algebra are subsets of Ω, where $B \subseteq \Omega$ corresponds to the event that the outcome of measurement is an element of B.

We say a POVM $M : \sigma(\Omega) \to \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ has global/local symmetry with respect to the group $H \subseteq \mathrm{U}(d)$ if for any $B \in \sigma(\Omega)$, the operator $M(B)$ has global/local symmetry with respect to H, i.e. it satisfies Eq.(1.1) or Eq.(1.2) respectively. Again, typically the local symmetry condition on a measurement is a much more restrictive condition.

In the following we first explore the consequences of a measurement having local symmetry and we see how in the case of gauge symmetries using the generalization of Schur-Weyl duality and in particular theorem 8, one can promote global symmetry of a measurement to a local symmetry (for states whose support is restricted to the symmetric or anti-symmetric subspace). Since the locally symmetric measurements typically are a much smaller class of measurements, this technique will be particularly useful in quantum estimation problems where one seeks to find the measurement that optimizes some figure of merit. Also, this trick is useful for determining whether a given estimation problem requires a nonlocal measurement on the n subsystems (i.e one that requires a quantum channel or entanglement) or whether a local measurement suffices. More generally, it can set an upper bound on the amount of entanglement required to achieve a particular degree of success in estimation. In the following we explain more about this.

One way to understand the restriction of local symmetry of measurement is via the following observation: Let the subgroup H of $\mathrm{U}(d)$ be a subgroup with unique Haar measure $d\mu$ and consider the twirling superoperator defined in Eq.(3.4). Then local symmetry of POVM $M : \sigma(\Omega) \to \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ with respect to H implies that $\mathcal{T}_H^{\otimes n}(M) = M$. This in turn implies that for any event $B \in \sigma(\Omega)$ and for any arbitrary density operator in $\mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ it holds that

$$\begin{aligned}\Pr(B) &= \mathrm{tr}\left(M(B)\rho\right) \\ &= \mathrm{tr}\left(\mathcal{T}_H^{\otimes n}\left(M(B)\right)\rho\right) = \mathrm{tr}\left(M(B)\mathcal{T}_H^{\otimes n}(\rho)\right)\end{aligned}$$

In other words, for any arbitrary state ρ if before measurement M, we apply the local twirling operation $\mathcal{T}_H$, then we do not disturb the statistics of the measurement M. Note that by applying the twirling operation before the measurement, we are mapping the state to $\mathrm{Alg}\{H'\}^{\otimes n}$ which typically can be much smaller than the space of all density operators in $\mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$. Applying this twirling operation decreases the size of the subsystems of the Hilbert space on which the state could be non-trivial and, as we will see later, this fact can set an upper bound on the amount of entanglement required to achieve a particular inference.

This is more clear in the case of gauge groups. Let $G_{\mathcal{A}}$ be the gauge group of a von Neumann algebra $\mathcal{A} \subseteq \mathrm{End}(\mathbb{C}^d)$. Then for any state ρ, the state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$ is in $\mathcal{A}^{\otimes n}$. Using the decomposition of the matrix algebra $\mathcal{A}$ given by Eq.(2.3), one can find a simple characterization of the form of state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$ for arbitrary ρ.

For instance, consider the Hilbert space $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R$ where $\mathcal{H}_L$ and $\mathcal{H}_R$ are two finite-dimensional Hilbert spaces. The system of interest decomposed into two subsystems: the left subsystem, described by $\mathcal{H}_L$, and the right subsystem, described by $\mathcal{H}_R$. Let the von Neumann Algebra $\mathcal{A}$ be $\mathrm{End}(\mathcal{H}_L) \otimes \mathbb{I}_{\mathcal{H}_R}$ where $\mathbb{I}_{\mathcal{H}_R} \in \mathrm{End}(\mathcal{H}_R)$ is the identity operator on $\mathcal{H}_R$. As we have seen in the above, for any measurement with local symmetry with respect to the group $G_{\mathcal{A}}$ the statistics of outcomes of the measurement on state ρ is exactly the same as the statistics of the outcomes of that measurement on state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$. But for any state $\rho \in \mathrm{End}\left(\mathcal{H}^{\otimes n}\right)$, it holds that $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho) \in \mathcal{A}^{\otimes n} = \mathrm{End}(\mathcal{H}_L^{\otimes n}) \otimes \mathbb{I}_{\mathcal{H}_R}^{\otimes n}$. In other words, this means that if before a measurement M with local symmetry with respect to $G_{\mathcal{A}}$, we discard all the n right subsystems, we still can simulate the measurement M by performing a measurement on the left subsystems. So, effectively the Hilbert space which is relevant in this problem is $\mathcal{H}_L$ which is of a smaller size than the Hilbert space $\mathcal{H}$. This clearly put an upper bound on the amount of entanglement required to implement measurement M. We can extend this argument to the case of an arbitrary von Neumann algebra $\mathcal{A}$.

A particularly important case is where $\mathcal{A}$ is a commutative algebra. In this case, for any arbitrary state $\rho \in \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$, the state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$, as an element of $\mathcal{A}^{\otimes n}$, commutes with all generators of $\mathcal{A}^{\otimes n}$. So if on each individual subsystem we measure an observable (projective von-Neumann measurement) inside the algebra $\mathcal{A}$ we will not change the state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$. But since $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho) \in \mathcal{A}^{\otimes n}$, we can uniquely specify $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$ by measuring a set of observables in $\mathcal{A}$ which generates the algebra $\mathcal{A}$ on each individual subsystem (note that generators of $\mathcal{A}$ all commute with each other and so can be measured simultaneously). However, after these measurements we know the exact description of state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$ and so we can then simulate any other measurement by a post-processing of the data we have gathered in these measurements. Finally, we notice that measuring generators of $\mathcal{A}$ on each individual subsystem for state $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$ gives exactly the same statistics as measuring these generators on the original state ρ. So we can summarize this discussion as follows.

Proposition 9 ***(Commutative Algebras)*** *Let $G_{\mathcal{A}}$ be the gauge group of the commutative von Neumann algebra $\mathcal{A} \subseteq End(\mathbb{C}^d)$. Then any measurement on $(\mathbb{C}^d)^{\otimes n}$ which has local symmetry with respect to $G_{\mathcal{A}}$ can be realized by*

measuring a set of observables which generate $\mathcal{A}$ on each system individually followed by a classical processing of the outcomes.

Therefore to implement a measurement which has local symmetry with respect to the gauge group $G_{\mathcal{A}}$ of a commutative algebra $\mathcal{A}$ one does not need any entanglement or adaptive measurement.

Having studied the consequences of local symmetry for measurements, we now show how the result of previous section and in particular theorem 8 implies that for states whose support is restricted to the symmetric/anti-symmetric subspace, the global symmetry of a measurement with respect to a gauge group can be promoted to a local symmetry.

Corollary 10 *(**Symmetry of Measurements**) Let $G_{\mathcal{A}}$ be the gauge group of a von Neumann algebra $\mathcal{A} \subseteq End(\mathbb{C}^d)$. Then for any POVM $M : \sigma(\Omega) \to End((\mathbb{C}^d)^{\otimes n})$ which has global symmetry with respect to $G_{\mathcal{A}}$ there is a POVM with local symmetry with respect to $G_{\mathcal{A}}$ (i.e. $\tilde{M} : \sigma(\Omega) \to \mathcal{A}^{\otimes n}$) which has exactly the same statistics for all states whose supports are confined to the symmetric (anti-symmetric) subspace. In particular, one can choose $\tilde{M}_{\pm} = \mathcal{L}_{\pm}(M)$ where $\mathcal{L}_{\pm}$ is the superoperator defined in theorem 8.*

Proof. First, recall that if $N : \sigma(\Omega) \to \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ is a POVM and $\mathcal{E}$ is a unital, positive quantum operation from $\mathrm{End}((\mathbb{C}^d)^{\otimes n})$ to itself, then $\mathcal{E}(N) : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ is also a POVM. By theorem 8 we know that $\mathcal{L}_{\pm}$ is a unital, completely positive map from $\mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ to itself. So $\tilde{M}_{\pm} \equiv \mathcal{L}_{\pm}(M)$ where $\tilde{M}_{\pm} : \sigma(\Omega) \to \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$ is also a POVM. Furthermore, theorem 8 implies that the image of $\mathcal{L}_{\pm}$ has local symmetry with respect to $G_{\mathcal{A}}$ (i.e. it is in $\mathcal{A}^{\otimes n}$). Finally, by definition, if POVM M has global symmetry with respect to $G_{\mathcal{A}}$ then for any $B \in \sigma(\Omega)$, $M(B)$ commutes with $\mathbf{Q}(G_{\mathcal{A}})$. Now since all elements of $\mathbf{Q}(G_{\mathcal{A}})$ are permutationally invariant they are block diagonal in irreps of the permutation group and in particular they commute with $\Pi_{\pm}$. So if $M(B)$ commutes with $\mathbf{Q}(G_{\mathcal{A}})$, then $\Pi_{\pm}M(B)\Pi_{\pm}$ will also commute with $\mathbf{Q}(G_{\mathcal{A}})$. Then using theorem 8 and the definition of $\tilde{M}_{\pm}$ we conclude that

$$\Pi_{\pm}\tilde{M}_{\pm}(B)\Pi_{\pm} = \Pi_{\pm}M(B)\Pi_{\pm} \tag{4.1}$$

Now consider the probability of event $B \in \sigma(\Omega)$ in the measurement described by POVM $\tilde{M}$ and state $\rho \in \mathrm{End}\left((\mathbb{C}^d)^{\otimes n}\right)$. This probability is given by $\Pr(B) = \mathrm{tr}(\rho\tilde{M}(B))$. Now if the support of ρ is restricted to the symmetric/anti-symmetric subspace then $\rho = \Pi_{\pm}\rho\Pi_{\pm}$ and so

$$\forall \mu: \quad \Pr(B) = \mathrm{tr}(\rho\tilde{M}(B)) = \mathrm{tr}\left(\rho\Pi_{\pm}\tilde{M}(B)\Pi_{\pm}\right)$$

Substituting Eq.(4.1) into this we conclude that

$$\begin{aligned}\Pr(B) &= \mathrm{tr}\left(\rho\Pi_{\pm}\tilde{M}(B)\Pi_{\pm}\right) \\ &= \mathrm{tr}\left(\rho\Pi_{\pm}M(B)\Pi_{\pm}\right) = \mathrm{tr}(\rho M(B))\end{aligned}$$

But $\mathrm{tr}(\rho M(B))$ is the probability of event B in the measurement described by POVM M performed on state ρ. Therefore measurement $\tilde{M}$ simulates measurement M. ■

Corollary 10 implies that if the support of state ρ is restricted to the symmetric/anti-symmetric subspace then any measurement with global symmetry with respect to $G_{\mathcal{A}}$ on ρ can be simulated by a measurement on $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$. In other words, if one is under the restriction of using measurements which have global symmetry with respect to $G_{\mathcal{A}}$ then by applying the channel $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}$ to a state which is restricted to the symmetric/anti-symmetric subspace does not lose any information. Note that generally the support of $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\rho)$ is not restricted to the symmetric(anti-symmetric) subspace.

Based on this observation one can put a strong condition on the form of measurements which can be useful, for instance, in finding the optimal measurement in a multi-copy estimation procedure (as we do in the next section). Note that for any given measurement with a global symmetry $G_{\mathcal{A}}$ there are many different other measurements which will have exactly the same statistics on all states whose support are restricted to the symmetric/anti-symmetric subspaces. These measurements may require different amounts of entanglement to be implemented. Having local symmetry with respect to $G_{\mathcal{A}}$ puts an upper bound on the amount of entanglement required to realize a measurement.

In particular, note that the combination of proposition 9 and corollary 10 implies that if a measurement has global symmetry with respect to $G_{\mathcal{A}}$ the gauge group of a commutative algebra $\mathcal{A}$ then among all possible measurements which can simulate this measurements on states with support in symmetric/anti-symmetric subspace there is one which does not need any entanglement to be realized.

Example

It is useful to consider a concrete example of the simulation of a measurement with global symmetry by one with local symmetry. To this end, consider a pair of qudits with the total Hilbert space $(\mathbb{C}^d)^{\otimes 2}$ and consider the unitary group of phase shifts $H_d \equiv \{e^{i\phi N} : \phi \in (0, 2\pi]\}$ where $N|i\rangle = i|i\rangle$ and $\{|i\rangle : i = 0 \cdots d-1\}$ is an orthonormal basis for $\mathbb{C}^d$. Note that the unitary group H_d is indeed a representation of U(1) on $\mathbb{C}^d$.

Now one can easily see that a measurement which has global (local) symmetry with respect to H_d has also global (local) symmetry with respect to $\{e^{i\phi_0}e^{i\phi N} : \phi_0, \phi \in (0, 2\pi]\}$ and vice versa. But in the specific case of $d = 2$, the latter group is a gauge group, as we have seen in section III A. In the case of $d = 2$ we denote $\{e^{i\phi_0}e^{i\phi N} : \phi_0, \phi \in (0, 2\pi]\}$ by G.

So, in the case of $d = 2$ according to corollary 10, we can infer that for states in the symmetric and antisymmetric subspaces, every measurement with global symmetry with respect to G (or equivalently with respect to

H_2) can be simulated with one that has local symmetry with respect to G (or equivalently with respect to H_2).

The measurements that have local symmetry are those for which all the POVM elements are *locally* diagonal in the eigenspaces of N, that is, in the basis $\{|0\rangle, |1\rangle\}$. All such measurements can be realized by a measurement of the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ followed by a classical post-processing of the outcome. Note that measurement in basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ can be realized by measuring N individually on each qubit. This is expected from proposition 9 because the algebra of commutants of the gauge group, is the algebra of diagonal matrices in the basis $\{|0\rangle, |1\rangle\}$ which is a commutative algebra.

On the other hand, POVM elements of any measurements that have global symmetry with respect to H_2 (or equivalently with respect to G) are those which commute with total number operator $N \otimes I + I \otimes N$ and so are block-diagonal relative to the eigenspaces of $N \otimes I + I \otimes N$. Let $M : \sigma(\Omega) \to \mathrm{End}\left((\mathbb{C}^2)^{\otimes 2}\right)$ be POVM of such measurement. Then for any $B \in \sigma(\Omega)$ it holds that

$$\Pi_{00} M(B) \Pi_{00} + \Pi_{11} M(B) \Pi_{11} + \Pi_{01} M(B) \Pi_{01} = M(B)$$

where $\Pi_{00} \equiv |00\rangle\langle 00|$, $\Pi_{11} \equiv |11\rangle\langle 11|$ and $\Pi_{01} \equiv |01\rangle\langle 01| + |10\rangle\langle 10|$. Therefore probability of event B for arbitrary state ρ is equal to

$$\begin{aligned}\mathrm{tr}\left(M(B)\rho\right) = &\,\mathrm{tr}\left(\Pi_{00}\rho\right)\mathrm{tr}\left(M(B)\Pi_{00}\right) \\ &+ \mathrm{tr}\left(\Pi_{11}\rho\right)\mathrm{tr}(M(B)\Pi_{11}) + \mathrm{tr}\left(\rho\Pi_{01}M(B)\Pi_{01}\right)\end{aligned}$$

Now if the state ρ is promised to be in the symmetric subspace, i.e. $\Pi_+ \rho \Pi_+ = \rho$ then

$$\begin{aligned}\mathrm{tr}\left(\rho\Pi_{01}M(B)\Pi_{01}\right) &= \mathrm{tr}\left(\left[\Pi_+\rho\Pi_+\right]\Pi_{01}M(B)\Pi_{01}\right) \\ &= \mathrm{tr}(\rho\Pi_{01})\mathrm{tr}\left(M(B)|\phi^+\rangle\langle\phi^+|\right)\end{aligned}$$

where $|\phi^+\rangle \equiv (1/\sqrt{2})(|01\rangle + |10\rangle)$. In other words,

$$\begin{aligned}\mathrm{tr}(M(B)\rho) =&\mathrm{Pr}(B|00)\mathrm{tr}(\Pi_{00}\rho) \\ &+ \mathrm{Pr}(B|11)\mathrm{tr}(\Pi_{11}\rho) + \mathrm{Pr}(B|01)\mathrm{tr}(\rho\Pi_{01})\end{aligned}$$

where

$$\begin{aligned}\mathrm{Pr}(B|00) \equiv \mathrm{tr}\left(M(B)\Pi_{00}\right), \;\; &\mathrm{Pr}(B|11) \equiv \mathrm{tr}\left(M(B)\Pi_{11}\right) \\ &\text{and} \;\; \mathrm{Pr}(B|01) \equiv \mathrm{tr}\left(M(B)|\phi^+\rangle\langle\phi^+|\right)\end{aligned}$$

and they can be interpreted as the conditional probability of event $B \in \sigma(\Omega)$ given each of the four outcomes. This means that to simulate this measurement one can measure N individually on each qubit, and based on the outcomes of these measurements choose outcome $\omega \in \Omega$ consistent with these conditional probabilities.

In other words, although the set of measurements with global symmetry is much larger than the set of measurements with local symmetry, all the information we can extract using a measurement with global symmetry can also be obtained by a measurement with local symmetry. Note that the measurement with local symmetry which we built based on the original measurement is exactly the same measurement as we can get by applying the superoperator $\mathcal{L}_+$ defined in theorem 8 to POVM of the original measurement. Also, note that from corollary 10 we know that this result holds for any arbitrary number of qubits.

In this example one can easily see that though implementing the measurement with global symmetry may require entanglement, implementing the measurement with local symmetry does not, nor does it require communication among the subsystems.

Finally, based on this example we provide another concrete instance that illustrates how the gauge property of the symmetry group is critical for being able to promote global symmetries to local symmetries. Consider the above example for the case of of $d = 3$ i.e. for qu*trits* rather than qubits.

The measurements that have local symmetry are those which can be obtained by classical post-processing of a measurement of the product basis $\{|ij\rangle : i, j = 0, 1, 2\}$, while those with global symmetry are merely block-diagonal with respect to the eigenspaces of total N. In particular, a measurement with global symmetry may include the rank-1 projectors onto the vectors $|11\rangle + |02\rangle + |20\rangle$ and $|11\rangle - (|02\rangle + |20\rangle)$, which both lie in the symmetric subspace. Such a measurement cannot be simulated by any measurement with local symmetry, which necessarily is unable to detect coherence between $|11\rangle$ and the subspace spanned by $|02\rangle + |20\rangle$.

V. MULTI-COPY ESTIMATION AND DECISION PROBLEMS

The main application of the duality is to multi-copy estimation problems. We begin by setting up a general framework for such problems.

Suppose Alice randomly chooses a state ρ from the density operators in $\mathrm{End}(\mathbb{C}^d)$ according to the probability density function p and then prepares n copies of this state and sends them to Bob through a quantum channel $\mathcal{E} : \mathrm{End}((\mathbb{C}^d)^{\otimes n}) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$. Bob's goal is to estimate some parameter(s) of state ρ. (We here adopt the convention that the term "estimation problem" includes decision problems as a special case). So upon receiving n systems he performs a measurement and generates some outcome in the outcome space Ω where Ω is a measure space, i.e. a set equipped with a σ-algebra $\sigma(\Omega)$ of subsets. The elements of the σ-algebra are subsets of Ω, where $B \subseteq \Omega$ corresponds to the event that Bob's measurement outcome is an element of B. The outcome space Ω can be continuous (in the case of general estimation problems) or discrete (in the case of decision problems).

In an arbitrary estimation strategy, Bob measures the n systems he has received and possibly does some post-processing on the outcome, ultimately generating an output in the set Ω. The entire strategy, which combines the measurement and the data processing, can be described by a POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$. For simplic-

ity, we will often refers to the estimation strategy as the measurement.

Therefore, the most general figure of merit which evaluates the performance of different strategies in an estimation problem is a function which assigns real numbers to all POVMs $M : \sigma(\Omega) \to \text{End}((\mathbb{C}^d)^{\otimes n})$. Equivalently, in the case of the multi-copy estimation problems we are considering here, the most general figure of merit can be described as a real functional which acts on the two variable function

$$q_M(B|\rho) = \text{tr}\left(M(B)\mathcal{E}(\rho^{\otimes n})\right),$$

the conditional probability that, using the strategy described by POVM $M : \sigma(\Omega) \to \text{End}((\mathbb{C}^d)^{\otimes n})$, the event $B \in \sigma(\Omega)$ happens given that Alice has chosen the state $\rho \in \text{supp}(p)$ and has sent state $\rho^{\otimes n}$ to Bob through the channel $\mathcal{E}$ (here, $\text{supp}(p)$ denotes the support of the distribution p).

This describes the most general figure of merit one can define for the multi-copy estimation problems we are considering here. However, in the particular cases where for example the goal is to estimate some parameter of ρ, say the expectation value of some observable for state ρ, one might use a figure of merit which only depends on the conditional probability of outcomes for different values of that parameter. Here, we think of the parameter as a random variable defined as a function of the state Alice chooses each time (The state is random and so any function of the state can be thought of as a random variable). Let $\mathfrak{s} : \text{supp}(p) \to \mathbb{R}$ be an arbitrary function from states in $\text{supp}(p)$ to real numbers. Then this function will map the random state ρ chosen by Alice to a random real variable $S = \mathfrak{s}(\rho)$. Then if Bob's goal is to estimate the value of parameter $\mathfrak{s}(\rho)$ for the state ρ which Alice has chosen each time (or to make a decision based on the value of this parameter) a reasonable family of figures of merit to evaluate Bob's performance can be expressed as functionals of

$$q_M(B|S \in \Delta),$$

where Δ is an interval of $\mathbb{R}$. This is the conditional probability that, using the strategy described by POVM $M : \sigma(\Omega) \to \text{End}((\mathbb{C}^d)^{\otimes n})$, event B happens given that the value of the random variable S is in Δ.

On the other hand, one can imagine the situations where, for example, the cost for wrong estimation of a parameter $S^{(1)}$ not only depends on the estimated value of $S^{(1)}$ and its actual value but also depends on the value of some other parameter say $S^{(2)}$ where $S^{(2)}$ is the random variable induced by the function $\mathfrak{s}^{(2)} : \text{supp}(p) \to \mathbb{R}$ acting on the random state Alice chooses. For instance, one may imagine situations where the cost of wrong estimation of a parameter $S^{(1)}$ depends also on the energy of state $\text{tr}(\rho H)$ where H is the Hamiltonian. So in this case $\mathfrak{s}^{(2)}(\cdot) = \text{tr}\left(H(\cdot)\right)$ defines a relevant parameter to evaluate the performance of the estimation procedure. In general, let $\vec{\mathfrak{s}}(\cdot) = \left(\mathfrak{s}^{(1)}(\cdot), \cdots, \mathfrak{s}^{(l)}(\cdot)\right)$ be a set of functions where each $\mathfrak{s}^{(i)}(\cdot)$ is a function from $\text{supp}(p)$ to $\mathbb{R}$. Then based on the set of functions $\vec{\mathfrak{s}}(\cdot) = \left(\mathfrak{s}^{(1)}(\cdot), \cdots, \mathfrak{s}^{(l)}(\cdot)\right)$ we can define a set of random variables $\left(S^{(1)}, \cdots, S^{(l)}\right)$ where the random variable $S^{(i)}$ is $\mathfrak{s}^{(i)}(\rho)$ where ρ is the random state Alice has chosen at each round. So a general figure of merit can be expressed as a functional of

$$q_M(B|\vec{S} \in \vec{\Delta}),$$

where $\vec{\Delta}$ is an l-dimensional interval of $\mathbb{R}^l$. This is the conditional probability that with Bob's strategy described by POVM $M : \sigma(\Omega) \to \text{End}((\mathbb{C}^d)^{\otimes n})$ event B happens given the value of the random variables $\vec{S}$ are in $\vec{\Delta}$.

The other reason to consider $q_M(B|\vec{S} \in \vec{\Delta})$ for more than one parameter $S^{(i)}$ is to study the cases where Bob is interested in estimating more than one parameter of the state.

Note that by having larger number of parameters l we can describe more and more general types of figure of merit. In general, if d is the dimension of $\mathbb{C}^d$ then the set of all (normalized) density operators can be specified by $d^2 - 1$ parameters. So having $l = d^2 - 1$ parameters is sufficient to specify the exact density operator Alice has chosen each time, and so $l = d^2 - 1$ parameters are sufficient to describe the most general form of figures of merit one can imagine for this problem (one can think of matrix elements of a density operator in a particular basis as different parameters). However, generally, having a figure of merit which can be defined using less than $d^2 - 1$ parameters, makes it easier to find the optimal estimation procedure.

To summarize, in the multi-copy estimation problem we are considering here, $q_M(B|\rho)$ has the maximal information required to evaluate the figure of merit of the strategy described by the POVM M. In other words, if for two different strateges described by POVMs $M : \sigma(\Omega) \to \text{End}((\mathbb{C}^d)^{\otimes n})$ and $M' : \sigma(\Omega) \to \text{End}((\mathbb{C}^d)^{\otimes n})$ it holds that

$$q_M\left(B|\rho\right) = q_{M'}\left(B|\rho\right) \tag{5.1}$$

for all $B \in \sigma(\Omega)$ and $\rho \in \text{supp}(p)$ then they will have exactly the same performance in the estimation problem with respect to any figure of merit. On the other hand, $q_M(B|\vec{S} \in \vec{\Delta})$ has generally less information i.e. it can be obtained by a coarse-graining of $q_M\left(B|\rho\right)$ but not necessarily vice versa. However, in many reasonable figures of merit one does not need to specify $q_M\left(B|\rho\right)$ to specify the figure of merit of the measurement M; it is sufficient to specify $q_M(B|\vec{S} \in \vec{\Delta})$. If this is the case, then even if Eq. (5.1) doesn't hold, as long as the weaker constraint

$$q_M\left(B|\vec{S} \in \vec{\Delta}\right) = q_{M'}\left(B|\vec{S} \in \vec{\Delta}\right) \tag{5.2}$$

holds for all $B \in \sigma(\Omega)$ and for all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability, then the two strategies yield the same performance for the figure of

merit of interest. Eq. (5.2) states that learning the outcome of measurement M is *precisely as informative about the parameter* $\vec{S}$ as learning the outcome of measurement M'.

Some examples of common figures of merit will be provided in Appendix D.

A. Main result

Scenario: Suppose that Alice randomly chooses an unknown state ρ from the density operators in $\mathrm{End}(\mathbb{C}^d)$ according to some probability density p (which we call the *single-copy prior*) and send n systems each prepared in the state ρ to Bob through a quantum channel $\mathcal{E} : \mathrm{End}((\mathbb{C}^d)^{\otimes n}) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$. Suppose that Bob makes measurements on the collection of n systems.

Let parameters $\vec{\mathfrak{s}}(\cdot) = \left(\mathfrak{s}^{(1)}(\cdot), \cdots, \mathfrak{s}^{(l)}(\cdot)\right)$ be an arbitrary set of functions where $\mathfrak{s}^{(i)} : \mathrm{supp}(p) \to \mathbb{R}$, and let $\vec{S}$ be the random variables defined as $\vec{S} \equiv \vec{\mathfrak{s}}(\rho)$ where ρ is the random state Alice chooses. We refer to $\vec{\mathfrak{s}}$ as the *parameters*. We say that the prior p is invariant under a subgroup H of $\mathrm{U}(d)$, or equivalently, *has H as a symmetry* if for all ρ we have

$$\forall V \in H : p(\rho) = p\left(V\rho V^\dagger\right). \tag{5.3}$$

We say that the parameter $\mathfrak{s}$ is invariant under a subgroup H of $\mathrm{U}(d)$, or equivalently, *has H as a symmetry* if for all $\rho \in \mathrm{supp}(p)$, i.e. all ρ assigned non-zero probability by the prior, we have

$$\forall V \in H : \vec{\mathfrak{s}}(\rho) = \vec{\mathfrak{s}}\left(V\rho V^\dagger\right). \tag{5.4}$$

We now present our main results, leaving the proofs to be presented in Sec. V D. We begin with a version of the result where the assumptions are particularly simple. These assumptions will be generalized shortly.

Theorem 11 *Let $\mathcal{A} \subseteq End(\mathbb{C}^d)$ be a von Neumann algebra, and let $G_{\mathcal{A}}$ be the gauge group associated with it. Then assuming that:*

1. *The prior p and the vector of parameters $\vec{\mathfrak{s}}$ have the gauge group $G_{\mathcal{A}}$ as a symmetry, and*
2. *The channel $\mathcal{E}$ is the identity channel, and*
3. *The prior p has support only on the pure states.*

then for any given measurement with POVM $M : \sigma(\Omega) \to End((\mathbb{C}^d)^{\otimes n})$, there is another measurement with POVM $M' : \sigma(\Omega) \to End((\mathbb{C}^d)^{\otimes n})$ whose image is entirely confined to $\mathcal{A}^{\otimes n}$ (i.e., $M' : \sigma(\Omega) \to \mathcal{A}^{\otimes n}$), such that M' is as informative about $\vec{S}$ as M is, i.e.,

$$q_M\left(B|\vec{S} \in \vec{\Delta}\right) = q_{M'}\left(B|\vec{S} \in \vec{\Delta}\right) \tag{5.5}$$

for all $B \in \sigma(\Omega)$ and all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability.

Remark 12 *An instance of the measurement described in theorem 11 is $M' \equiv \mathcal{L}_+(M)$, where $\mathcal{L}_+$ is the unital quantum channel defined in Eq.* (3.10)*.*

One can generalize this theorem in two ways: from the identity channel to a class of nontrivial channels, and from a prior that has support only on pure states to a certain class of priors that have support on mixed states. We begin by defining the classes in question.

We define a channel $\mathcal{E}$ to be *noiseless on $\mathcal{A}^{\otimes n}$* if for all states ρ in $\mathrm{End}((\mathbb{C}^d)^{\otimes n})$, $\mathcal{E}(\rho)$ and ρ have the same reduction on the algebra $\mathcal{A}^{\otimes n}$, i.e.,

$$\forall R \in \mathcal{A}^{\otimes n} : \mathrm{tr}(R\mathcal{E}(\rho)) = \mathrm{tr}(R\rho), \tag{5.6}$$

or equivalently, $\mathcal{T}^{\otimes n}_{G_{\mathcal{A}}} \circ \mathcal{E} = \mathcal{T}^{\otimes n}_{G_{\mathcal{A}}}$.

Let prior $\tilde{p}$ be one with support confined to the pure states. Define a prior p to be a *$G_{\mathcal{A}}$-distortion* of $\tilde{p}$ if it can be realized by sampling a pure state from $\tilde{p}$ and then applying a quantum channel $\mathcal{N} : \mathrm{End}(\mathbb{C}^d) \to \mathrm{End}(\mathbb{C}^d)$ to the state, where $\mathcal{N}$ is noiseless on $\mathcal{A}$ and is also $G_{\mathcal{A}}$-covariant i.e. $\forall V \in G_{\mathcal{A}} : \mathcal{N}(\cdot) = \mathcal{N}(V \cdot V^\dagger)$. We then have the following generalization of the theorem 11.

Theorem 13 *(Generalization of theorem 11) The implication in theorem 11 still holds if one weakens assumptions 2 and 3 to*
2′. The channel $\mathcal{E}$ is noiseless on $\mathcal{A}^{\otimes n}$.
3′. The prior p is a $G_{\mathcal{A}}$-distortion of one that has support only on the pure states.

Remark 14 *An instance of the measurement described in theorem 13 is $M' \equiv \mathcal{L}_+ \circ (\mathcal{N}^\dagger)^{\otimes n} \circ \mathcal{E}^\dagger(M)$, where $\mathcal{L}_+$ is the unital quantum channel defined in Eq.* (3.10)*.*

We now make explicit what our main theorem implies for multi-copy estimation problems.

Corollary 15 *If the figure of merit for a strategy M in the n-copy estimation problem can be expressed as a functional of $q_M(B|\vec{S} \in \vec{\Delta})$ for some set of parameters $\vec{\mathfrak{s}}$, then if the assumptions of the theorem 13 (or theorem 11) hold for an algebra $\mathcal{A}$, it follows that the POVM elements of the optimal measurement can be chosen to be in $\mathcal{A}^{\otimes n}$.*

Corollary 15 implies that the optimal measurement has the gauge group $\mathrm{G}_{\mathcal{A}}$ as a local symmetry. Then, in the special case wherein the algebra $\mathcal{A}$ is commutative, by proposition 9, it follows that it can be implemented by measuring a set of observables which generates $\mathcal{A}$ separately on each of the n systems and then performing a classical processing on the outcomes.

To apply corollary 15, the figure of merit for an estimation strategy M must be a functional of the conditional $q_M(B|\vec{S} \in \vec{\Delta})$. In appendix D, we demonstrate that this is indeed the case for two of the most common figures of merit: the expected cost for an arbitrary cost function, and the mutual information between the estimated and true value of a parameter. So the condition on the figure of merit is generically satisfies.

B. The reduction of the state to the algebra

We here describe an alternative way to state assumption 1 of theorem 11 in the case where the prior p has support only on pure states.

We begin with a definition. We say that a function over states $g : \mathrm{End}(\mathbb{C}^d) \to \mathbb{R}$ *depends only on the reduction of the state to the algebra* $\mathcal{A}$ if it can be expressed as

$$g(\rho) = f\left(\mathrm{tr}(\rho\tilde{A}_1), \cdots, \mathrm{tr}(\rho\tilde{A}_D)\right)$$

for some function $f : \mathbb{C}^D \to \mathbb{C}$, where $\{\tilde{A}_1, \cdots, \tilde{A}_D\} \subset \mathcal{A}$ is a basis for $\mathcal{A}$.

In terms of this notion, the alternative statement of assumption 1 is:

1′. *The prior p and the vector of parameters* $\vec{\mathfrak{s}}$ *depend only the reduction of the state to the algebra* $\mathcal{A}$.

The fact that assumption 1′ implies assumption 1 is clear: If $V \in G_{\mathcal{A}}$ then $\mathrm{tr}(\rho V^\dagger A V) = \mathrm{tr}(\rho A)$ for arbitrary density operator ρ in $\mathrm{End}(\mathbb{C}^d)$ and arbitrary $A \in \mathcal{A}$. Then since according to assumption 1′, p and $\vec{\mathfrak{s}}$ can be expressed as a function of $\mathrm{tr}(\rho\tilde{A}_1), \cdots, \mathrm{tr}(\rho\tilde{A}_D)$ we conclude that $p(V\rho V^\dagger) = p(\rho)$ and $\vec{\mathfrak{s}}(V\rho V^\dagger) = \vec{\mathfrak{s}}(\rho)$ for arbitrary ρ and arbitrary $V \in G_{\mathcal{A}}$.

The fact that assumption 1 implies assumption 1′ is true because of the following: First, note that for any two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$, if $\langle\psi_1|\tilde{A}_i|\psi_1\rangle = \langle\psi_2|\tilde{A}_i|\psi_2\rangle$ for $\{\tilde{A}_1, \cdots, \tilde{A}_D\}$ then there exists $V \in G_{\mathcal{A}}$ such that $V|\psi_1\rangle = |\psi_2\rangle$ [6]. Second, according to assumption 1, $\vec{\mathfrak{s}}$ is invariant under $G_{\mathcal{A}}$ and so

$$\vec{\mathfrak{s}}(|\psi_2\rangle\langle\psi_2|) = \vec{\mathfrak{s}}\left(V|\psi_1\rangle\langle\psi_1|V^\dagger\right) = \vec{\mathfrak{s}}\left(|\psi_1\rangle\langle\psi_1|\right)$$

Therefore the value of $\vec{\mathfrak{s}}(|\psi_1\rangle\langle\psi_1|)$ (and similarly $p(|\psi_1\rangle\langle\psi_1|)$) is uniquely specified by $\langle\psi_1|\tilde{A}_i|\psi_1\rangle$ for $\{\tilde{A}_1, \cdots, \tilde{A}_D\}$ which completes the proof.

Note that the restriction to pure states plays an essential role in the equivalence of assumptions 1 and 1′ and this equivalence cannot be extended to the case of mixed states, i.e. in general a parameter $\mathfrak{s}$ which satisfies assumption 1, $\mathfrak{s}(\rho)$ cannot be expressed as a function of $\mathrm{tr}(\rho\tilde{A}_1), \ldots, \mathrm{tr}(\rho\tilde{A}_D)$ if ρ is mixed. For instance, consider the case where $\mathcal{A}$ is the trivial algebra generated by the identity operator, so that $G_{\mathcal{A}}$ is the group of all unitaries on $\mathbb{C}^d$. In this case, the identity operator is a basis for $\mathcal{A}$ and consequently every state ρ has the same reduction to $\mathcal{A}$. This means that the only functions that depend only on the reduction of the state to $\mathcal{A}$ are constant functions. However, there exist non-constant functions $\mathfrak{s}$, such as $\mathfrak{s}(\rho) = \mathrm{tr}(\rho^2)$, which are invariant under the group of all unitaries and therefore have the symmetry property required to satisfy assumption 1. So the equivalence of assumption 1 and assumption 1′ cannot be extended to the case of mixed states.

C. Examples

1. *Estimating parameters defined by a single observable*

Recall the problem considered by Hayashi *et al.* [3] and discussed in the introduction. The problem is to estimate the expectation value of a *single* observable A based on n copies of a state. Casting this in our language, the vector of parameters to be estimated, $\vec{\mathfrak{s}}(\rho)$, has only a single component, $\mathfrak{s}(\rho) = \mathrm{tr}(A\rho)$. The figure of merit considered in Ref. [3] is the expected cost where the cost function is the squared error, i.e.

$$C(s_{est}, \mathfrak{s}(\rho)) = (s_{est} - \mathfrak{s}(\rho))^2.$$

Finally, the prior they consider is the unitarily-invariant measure over pure states and the channel $\mathcal{E}$ between the source and the estimator is the identity channel. It follows that the assumptions of theorem 11 are all satisfied for the algebra $\mathcal{A} = \mathrm{Alg}\{A, I\}$. Furthermore, one can show that the squared error for a measurement M is a functional of the conditional $q_M(s_{est} \in \Delta_{est} | S \in \Delta)$ in which S is the actual value of the parameter, s_{est} is the estimated value, Δ and Δ_{est} are two arbitrary intervals in $\mathbb{R}$ (see Appendix D 1). So the assumptions of corollary 15 are satisfied. Consequently, the optimal measurement can be confined to $\mathcal{A}^{\otimes n}$, but given that $\mathcal{A}$ is commutative, it follows from corollary 9 that it can be implemented by measuring the observable A separately on each system and performing classical data processing on the outcomes. So we have shown that the result of Hayashi *et al.* is recovered as a special case of ours.

It is worth noting that for estimation problems involving only a single observable A (or a set of commuting observables, which amounts to the same), there is in fact a very broad class of problems for which the optimal estimation can be achieved by separate measurements of A on each system. Indeed, one can consider the estimation of any parameter that depends only on A, i.e. any function of the form $f(\mathrm{tr}(\rho A), \mathrm{tr}(\rho A^2), \mathrm{tr}(\rho^2 A^2), \ldots)$. This includes the estimation of higher order moments of A, decisions about the sign of the expectation value of A, etcetera. One can also take the prior p to be arbitrary over pure states as long as it depends only on A. Also prior p can be nonzero on mixed states as long as p is a $G_{\mathcal{A}}$-distortion of a prior which is nonzero only on pure states. Finally, there are many choices for the figure of merit. We mention only two. We could take the mutual information between the estimated values of the parameters and their actual values, or we could take the expected cost for an arbitrary cost function that depends only on A. For all of these cases, the figure of merit for an estimation strategy M is a functional of $q_M(B|\vec{S} \in \vec{\Delta})$ (see Sec. D 1), so as long as the prior p and the channel $\mathcal{E}$ satisfy assumptions 2′ and 3′ of theorem 13, all the assumptions of corollary 15 are satisfied, and separate measurements of A suffice. Our result therefore constitutes a very significant generalization of the previously know results.

2. Decision problem for a single qubit

Suppose we are given n copies of qubit state ρ, a density operator in $\mathrm{End}(\mathbb{C}^2)$. For $b \in 0, 1$, define

$$|\psi(\theta, b)\rangle \equiv \cos\alpha_b|0\rangle + e^{i\theta}\sin\alpha_b|1\rangle$$

where α_0 and α_1 are distinct angles in the range $[0, \pi)$ and where $\theta \in [0, 2\pi)$. Assume the single-copy prior $p(\rho)$ is as follows: the state is drawn from the set $\{|\psi(\theta, b)\rangle\}$ where θ is uniformly distributed over $[0, 2\pi)$ and b has uniform distribution over $\{0, 1\}$. This prior is illustrated in Fig. 1(a). The goal is to get information about the value of the bit b using n copies of state given according to this single-copy prior (this example is a decision problem). For instance, one might be interested to determine the value of the bit b with minimum probability of error. In general, we assume the goal is to generate an outcome in the outcome set Ω with σ-algebra $\sigma(\Omega)$ and the performance of different strategies are evaluated by a figure of merit which can be expressed as a functional acting on $q(B|b = b_0)$, i.e., the probability of event $B \in \sigma(\Omega)$ while the value of b is $b_0 \in \{0, 1\}$.

In this case, the parameter to be estimated is defined by $\mathfrak{s}(|\psi(\theta, b)\rangle\langle\psi(\theta, b)|) = b$. Adopting the convention that $|0\rangle$ and $|1\rangle$ are eigenstates of the Pauli observable σ_z, it is clear that the prior p and the parameter to be estimated, $\mathfrak{s}$, are both invariant under unitaries of the form $e^{i\phi'}e^{i\phi\sigma_z}$ where $\phi, \phi' \in [0, 2\pi)$, which describe phase shifts or rotations about the axis $\hat{z}$. As we have seen in the section III A this group is a gauge group. The algebra that corresponds to the commutant of this gauge group is $\mathcal{A} = \mathrm{Alg}\{\sigma_z, I\}$. Finally, since the figure of merit depends only on $q(B|b = b_0)$ the assumptions of corollary 15 are satisfied (Note that since $\mathfrak{s}(|\psi(\theta, b)\rangle\langle\psi(\theta, b)|) = b$, b can be thought as the random variable defined by parameter $\mathfrak{s}$ acting on states.). Therefore, we can infer that to achieve the optimal estimation, it suffices to consider POVMs inside the algebra $\mathcal{A}^{\otimes n}$ and since $\mathcal{A}$ is commutative, it suffices to measure σ_z on each system individually. In other words, all the information we can get from the state $|\psi(\theta, b)\rangle^{\otimes n}$ about the value of b we can also get from the mixed state $[\cos^2(\alpha_b)|0\rangle\langle 0| + \sin^2(\alpha_b)|1\rangle\langle 1|]^{\otimes n}$.

Note, however, that if one acquires some information about θ, then this information can be useful for estimating b: In the extreme case where we know the exact value of θ, we can perform the Helstrom measurement [11] for distinguishing the two pure states $|\psi(\theta, 0)\rangle^{\otimes n}$ and $|\psi(\theta, 1)\rangle^{\otimes n}$. So one estimation strategy is to use some of the qubits to estimate θ and then use this information to choose an optimal measurement for estimating b. But our result shows that by this strategy one cannot get more information than what one gets by ignoring θ and measuring σ_z on individual systems. [Note that this result also implies that to get information about θ from each system we necessarily disturb its information about b. This can be interpreted as an example of information-disturbance tradeoff.]

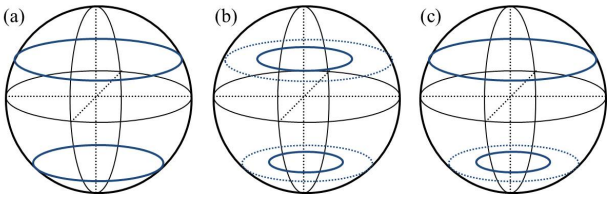

FIG. 1: The Bloch ball representation of the quantum states of a single qubit for three variations of a decision problem. The pair of circles in each case indicate the support of the single-copy prior over states and the goal is to decide which circle the state is drawn from, given n copies of the state. (a) A prior with support confined to the pure states. (b) A prior that is a gauge distortion of the first. (c) A prior for which unentangled measurement will not be generally sufficient to achieve optimal estimation.

Generalization to priors whose support is not confined to pure states. Theorem 13 implies that measuring σ_z on each system is optimal even if the single-copy prior is a $\mathrm{G}_{\mathcal{A}}$-distortion of the one described above. In this case a $\mathrm{G}_{\mathcal{A}}$-distortion is implemented by a channel $\mathcal{N}$ that is covariant under phase shifts and noiseless on $\mathrm{Alg}\{\sigma_z, I\}$. The only channels having these properties are those corresponding to dephasing about the $\hat{z}$ axis (i.e. $\mathcal{N}(\rho) = (1-r)\rho + r\,\sigma_z\rho\sigma_z$ for $0 < r < 1$). In the single-copy prior which is achieved by this distortion, the state is drawn from the set $\{\rho(\theta, b) \equiv \mathcal{N}(|\psi(\theta, b)\rangle\langle\psi(\theta, b)|)\}$ where b and θ are distributed as before (for a given b, this describes a circle *within* the Bloch ball). The parameter to be estimated is $\mathfrak{s}(\rho(\theta, b)) = b$. Note that both the prior and the parameter in this new estimation problem are invariant under the group of phase shifts. Corollary 15 implies that the estimation problem so defined is also one wherein the optimal estimation is achieved by implementing a measurement of σ_z on each qubit.

Example where unentangled measurements are generally not sufficient. Now suppose we are given n copies of state $\{\rho(\theta, b)\}$ where $\rho(\theta, 0) = |\psi(\theta, 0)\rangle\langle\psi(\theta, 0)|$ and $\rho(\theta, 1) = \mathcal{N}(|\psi(\theta, 1)\rangle\langle\psi(\theta, 1)|)$ where $\mathcal{N}$ is an arbitrary dephasing channel and where again θ is uniformly distributed between $(0, 2\pi]$ and b has arbitrary distribution. Effectively, we have a U(1)-orbit of pure states (a circle on the Bloch sphere) for $b = 0$, and a dephased version of a distinct U(1)-orbit (a circle within the Bloch ball) for $b = 1$. Again the goal is to find the value of the bit b.

This estimation problem satisfies assumption 1 of theorems 11 and 13 because the prior and the parameter have the same gauge group symmetry as the other examples considered in this section. We seek to show that nonetheless, in this case, the optimal measurement is *not* achieved by performing separate measurements of σ_z on each qubit.

To see this, note first that because the $b = 0$ states are pure while the $b = 1$ states are mixed, the purity of state does contain information about b. Now consider the projective measurement which projects state to the different

irreps of $\mathcal{S}_n$ which show up in the representation $\mathbf{P}(\mathcal{S}_n)$ on $(\mathbb{C}^d)^{\otimes n}$. It is well known that this von Neumann measurement is highly nonlocal and requires interaction between all n systems. This projective measurement is one that reveals information about the eigenvalues of the single-copy density operator and hence about its purity, as the following argument demonstrates [12].

First, note that if the single-copy state is pure, then the n-copy state is in the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$ and the outcome of the above projective measurement is fixed. On the other hand, if the single-copy state is mixed, then there is always a nonzero probability that the measurement projects the state to a subspace other than the symmetric subspace. In other words, there is a nonzero probability that the outcome of this measurement achieves an unambiguous discrimination between the mixed state case and the pure state case. This implies that there is a nonzero probability of determining the true value of b unambgiuously. However, one can easily see that for the given prior by measuring σ_z on each qubit it is not possible to unambigiously determine the true value of the bit b. Therefore, at least for some figures of merit, entangled measurements have advantage over unentangeled measurements.

Incidentally, note that since the state of the total n systems is a permutationally-invariant state, i.e. it commutes with $\mathbf{P}(\mathcal{S}_n)$ it is block diagonal in the irreps of $\mathcal{S}_n$ that show up in the representation of $\mathbf{P}(\mathcal{S}_n)$. Therefore by performing the von Neumann measurement which projects into these blocks, the final state (forgetting the outcome of this measurement) will be the same as the initial state and therefore the statistics of any subsequent measurement will not be affected, that is, implementing such a measurement does not compromise the informativeness of any other measurement.

This phenomenon is generic. In multi-copy decision problems in which the goal is to distinguish between a mixed state and a pure state, entangled measurements can achieve a better performance than unentangled measurements (at least with respect to some figures of merits).

3. Decision problem for pair of qubits

In the previous example we assumed a bit is encoded in the state of one qubit and the goal is to acquire information about that bit using n copies of that state. Now suppose we modify the example in the following way: We assume each system consists of two qubits (rather than one), left and right, i.e. the Hilbert space of each system is $\mathbb{C}^4 \cong \mathcal{H}_L \otimes \mathcal{H}_R$. Again, we are given n copies of state ρ according to the single-copy prior $p(\rho)$ which is defined as follows: the state is drawn from the set

$$\{(I \otimes V)|\psi(b)\rangle_{LR}\}$$

where b is uniformly distributes on $b \in \{0,1\}$, V is distributed according to the Haar measure over U(2) and

$$|\psi(0)\rangle_{LR} = |00\rangle_{LR} \ , \quad |\psi(1)\rangle_{LR} = \frac{|01\rangle_{LR} + |10\rangle_{LR}}{\sqrt{2}}.$$

The goal is again to get information about the bit b and therefore, the parameter to be estimated is defined implicitly by the condition that $\mathfrak{s}\left((I \otimes V)|\psi(b)\rangle_{LR}\right) = b$. It is then clear that the group of all unitaries acting on the right qubit, i.e. $\{I \otimes V\}$ where $V \in \mathrm{U}(2)$, is a symmetry group of both the prior p and the parameter $\mathfrak{s}$. Moreover this group of unitaries is clearly a gauge group, so it is a gauge symmetry of the prior and the parameter. The algebra associated with this gauge group is the full algebra of operators on the left qubit, i.e. $\mathcal{A} \equiv \mathrm{End}(\mathcal{H}_L) \otimes I$.

Again, we can see that for any figure of merit which depends only on $q(B|b = b_0)$, the assumptions of corollary 15 are satisfied and therefore to achieve the optimal estimation, it suffices to consider measurement operators inside the algebra $\mathcal{A}^{\otimes n}$. It follows that it suffices to consider measurements that are nontrivial on the left qubits only. In other words, one can essentially ignore the right qubits. However, note that deciding about the value of b is also equivalent to deciding that whether the reduced state of the right qubits is $(V|0\rangle)^{\otimes n}$ or $(I/2)^{\otimes n}$. It follows that the n right qubits do contain some information about the value of b, however, our results imply that once one has the information contained in the left qubits, the information contained in the right qubits is redundant.

D. Proof of theorem 11 and theorem 13

To prove theorem 11 we first prove the following lemma which holds for any arbitrary subgroup of the unitary group.

Lemma 16 ***(From symmetry of the problem to symmetry of the measurement)*** *In the scenario described in section V A, assume the prior p and the vector of parameters $\vec{\mathfrak{s}}$ are invariant under a subgroup H of $U(d)$ which has the (normalized) Haar measure $d\mu$. Then for any measurement described by a POVM $M : \sigma(\Omega) \to End((\mathbb{C}^d)^{\otimes n})$, the measurement described by*

$$\tilde{M} \equiv \mathcal{T}_{\boldsymbol{Q}(H)}(M) = \int_H d\mu(V)\ V^{\otimes n} M {V^\dagger}^{\otimes n}$$

is equally informative about $\vec{\mathfrak{s}}$, that is,

$$q_M\left(B|\vec{S} \in \vec{\Delta}\right) = q_{\tilde{M}}\left(B|\vec{S} \in \vec{\Delta}\right) \tag{5.7}$$

for all $B \in \sigma(\Omega)$ and all l-dimensional intervals $\vec{\Delta} \subseteq \mathbb{R}^l$ which are assigned nonzero probability.

Proof. First note that for any $B \in \sigma(\Omega)$

$$q_M(B|\rho) = \mathrm{tr}\left(\rho^{\otimes n} M(B)\right)$$

and

$$q_{\tilde{M}}(B|\rho) = \mathrm{tr}\left(\rho^{\otimes n}[\int_H d\mu(V)\ V^{\otimes n} M(B) V^{\dagger\otimes n}]\right)$$

Therefore, by the cyclic property of the trace,

$$q_{\tilde{M}}(B|\rho) = \int_H d\mu(V)\ q_M(B|V\rho V^\dagger) \tag{5.8}$$

On the other hand,

$$q_M\left(B|\vec{S}\in\vec{\Delta}\right) = \frac{1}{\Pr(\vec{S}\in\vec{\Delta})}\int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ \ q_M(B|\rho) \tag{5.9}$$

and similarly

$$q_{\tilde{M}}\left(B|\vec{S}\in\vec{\Delta}\right) = \frac{1}{\Pr(\vec{S}\in\vec{\Delta})}\int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ \ q_{\tilde{M}}(B|\rho) \tag{5.10}$$

where $\Pr(\vec{S}\in\vec{\Delta})$ is defined as

$$\Pr(\vec{S}\in\vec{\Delta}) \equiv \int_{\vec{\mathfrak{s}}(\rho)\in\vec{\Delta}} \mathrm{d}\rho\ p(\rho). \tag{5.11}$$

But

$$\begin{aligned}
&\int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ \ q_{\tilde{M}}(B|\rho)\\
&= \int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ \int_H d\mu(V)\ q_M(B|V\rho V^\dagger)\\
&= \int_H d\mu(V)\int_{\vec{\mathfrak{s}}(V\rho V^\dagger)\in\vec{\Delta}} d\rho\ p(V\rho V^\dagger)\ q_M(B|V\rho V^\dagger)\\
&= \int_H d\mu(V)\int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ \ q_M(B|\rho)\\
&= \int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ \ q_M(B|\rho)
\end{aligned}$$

where to get the second line we use Eq.(5.8), to get the third line we use the invariance of p and $\vec{\mathfrak{s}}$ under G, to get the fourth line we change the integral variable from ρ to $V\rho V^\dagger$ and to get the last line we use the fact that Haar measure is normalized. This completes the proof. ■

Proof. (**Theorem** 11)

According to the first condition in theorem 11, the prior p and the parameters $\vec{\mathfrak{s}}$ are invariant under the gauge group $G_{\mathcal{A}}$. So we can use lemma 16 for the symmetry group $G_{\mathcal{A}}$. This implies that for any given POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ and

$$\tilde{M} \equiv \mathcal{T}_{col}(M) = \int_{G_{\mathcal{A}}} d\mu(V)\ V^{\otimes n} M V^{\dagger\otimes n} \tag{5.12}$$

it holds that

$$q_{\tilde{M}}\left(B|\vec{S}\in\vec{\Delta}\right) = q_M\left(B|\vec{S}\in\vec{\Delta}\right) \tag{5.13}$$

for all $B \in \sigma(\Omega)$ and all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability. Now according to assumption 3 of theorem 11, the prior p is nonzero only for pure states. So for all states in $\{\rho^{\otimes n} : \rho \in \mathrm{supp}(p)\}$, i.e. the states Alice is sending to Bob, the support of the state is restricted to the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$. Since, by assumption 2, the channel is assumed to be the identity map, Bob receives the same state. Therefore all states that Bob receives are restricted to the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$. This, together with the fact that the measurement $\tilde{M}$ has global symmetry , and using corollary 10, implies that $\forall B \in \sigma(\Omega)$ and $\forall \rho \in \mathrm{supp}(p)$

$$\mathrm{tr}\left(\tilde{M}(B)\rho^{\otimes n}\right) = \mathrm{tr}\left(\mathcal{L}_+(\tilde{M}(B))\rho^{\otimes n}\right)$$

Define $M' \equiv \mathcal{L}_+(\tilde{M})$ where $\mathcal{L}_+$ is the superoperator defined in Eq.(3.10) of theorem 8. Then the above equality implies that

$$q_{\tilde{M}}\left(B|\vec{S}\in\vec{\Delta}\right) = q_{M'}\left(B|\vec{S}\in\vec{\Delta}\right) \tag{5.14}$$

for all $B \in \sigma(\Omega)$ and all $\vec{\Delta}$ which are assigned nonzero probability. This together with Eq.(5.13) implies that for arbitrary POVM M

$$q_M\left(B|\vec{S}\in\vec{\Delta}\right) = q_{M'}\left(B|\vec{S}\in\vec{\Delta}\right) \tag{5.15}$$

for all $B \in \sigma(\Omega)$ and all $\vec{\Delta}$ which are assigned nonzero probability.

Finally, using the fact that Π_+ commutes with $V^{\otimes n}$ for arbitrary $V \in \mathrm{U}(d)$ we can easily see that

$$\mathcal{L}_+(\tilde{M}) = \mathcal{L}_+(M),$$

so that

$$M' = \mathcal{L}_+(M).$$

From theorem 8, we know that the image of $\mathcal{L}_+$ is in $\mathcal{A}^{\otimes n}$ and therefore so is $M'(B)$ for arbitrary $B \in \sigma(\Omega)$. ■

Proof. (**Theorem 13**)

We first prove the special case of theorem 13 where assumptions 1, 2' and 3 hold. In other words, we first prove the theorem for the case of general channels which satisfy the assumptions of theorem 13 but for the special case where the prior is still nonzero only on pure states. Then we extend the result to the case of general priors which satisfy the assumption 3'.

(i) Generalization to non-identity channels, pure state priors:

The idea is to convert the estimation problem with channel $\mathcal{E}$ to another estimation problem with the identity channel and then apply the result of theorem 11 to this new estimation problem.

For any estimation problem described by the parameters $\vec{\mathfrak{s}}$, prior p, and the channel $\mathcal{E}$, we consider the two following scenarios:

- Scenario (a) in which Alice prepares n copies of the state ρ according to the probability density $p(\rho)$ and sends them through the channel $\mathcal{E}$ and then Bob performs a measurement described by POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$, and

- Scenario (b) in which Alice prepares n copies of the state ρ according to the probability density $p(\rho)$ but then sends them through the identity channel and Bob performs the measurement described by POVM $\mathcal{E}^\dagger(M)$ on the systems.

The definitions of these two scenarios immediately imply

$$q^{(a)}_M\left(B|\vec{S}\in\vec{\Delta}\right) = q^{(b)}_{\mathcal{E}^\dagger(M)}\left(B|\vec{S}\in\vec{\Delta}\right) \tag{5.16}$$

where the left and right hand sides describe the conditional for the scenarios (a) and (b) respectively. This is true because in the scenario (a) the probability of event $B \in \sigma(\Omega)$ given that Alice has chosen state ρ is $\mathrm{tr}\left(M(B)\mathcal{E}(\rho^{\otimes n})\right)$. On the other hand, in the scenario (b), the probability of event $B \in \sigma(\Omega)$ given that the state chosen by Alice is ρ is $\mathrm{tr}\left(\mathcal{E}^\dagger(M(B))\rho^{\otimes n}\right)$. But since

$$\mathrm{tr}\left(M(B)\mathcal{E}(\rho^{\otimes n})\right) = \mathrm{tr}\left(\mathcal{E}^\dagger(M(B))\rho^{\otimes n}\right)$$

for all $\rho \in \mathrm{supp}(p)$ and $B \in \sigma(\Omega)$, Eq.(5.16) follows.

Now in the scenario (b), where the channel is the identity map, we can apply theorem 11. Note that the assumptions of this theorem are satisfied for the gauge group $G_{\mathcal{A}}$. This implies

$$q^{(b)}_{\mathcal{L}_+(\mathcal{E}^\dagger(M))}\left(B|\vec{S}\in\vec{\Delta}\right) = q^{(b)}_{\mathcal{E}^\dagger(M)}\left(B|\vec{S}\in\vec{\Delta}\right) \tag{5.17}$$

Since the channel $\mathcal{E}$ is noiseless on $\mathcal{A}^{\otimes n}$ (assumption $2'$) then all elements of $\mathcal{A}^{\otimes n}$ are fixed points of $\mathcal{E}^\dagger$. (The fact that $\mathcal{E}$ is noiseless on $\mathcal{A}^{\otimes n}$ implies that for any operators $R_1 \in \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ and $R_2 \in \mathcal{A}^{\otimes n}$ it holds that $\mathrm{tr}\left(R_2\mathcal{E}(R_1)\right) = \mathrm{tr}\left(R_2R_1\right)$. But this implies that $\mathrm{tr}\left(\mathcal{E}^\dagger(R_2)R_1\right) = \mathrm{tr}\left(R_2R_1\right)$ which proves the claim.)

Then since elements of $\mathcal{A}^{\otimes n}$ are fixed points of $\mathcal{E}^\dagger$ and since the image of $\mathcal{L}_+$ is in $\mathcal{A}^{\otimes n}$ we conclude that

$$\mathcal{E}^\dagger \circ \mathcal{L}_+ = \mathcal{L}_+$$

Putting this into Eq.(5.17) we find

$$q^{(b)}_{\mathcal{E}^\dagger\circ\mathcal{L}_+\circ\mathcal{E}^\dagger(M)}\left(B|\vec{S}\in\vec{\Delta}\right) = q^{(b)}_{\mathcal{E}^\dagger(M)}\left(B|\vec{S}\in\vec{\Delta}\right)$$

Now for the conditionals on each side of this equality, we use Eq.(5.16) to find the measurement in the scenario (a) that yields the same conditional. We infer that

$$q^{(a)}_{\mathcal{L}_+\circ\mathcal{E}^\dagger(M)}\left(B|\vec{S}\in\vec{\Delta}\right) = q^{(a)}_M\left(B|\vec{S}\in\vec{\Delta}\right), \tag{5.18}$$

and this holds for arbitrary $\rho \in \mathrm{supp}(p)$ and event $B \in \sigma(\Omega)$ and arbitrary POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$. This completes the proof of the special case of the theorem where the prior p is nonzero only for pure states. Note that in this particular case one can choose

$$M' \equiv \mathcal{L}_+ \circ \mathcal{E}^\dagger(M).$$

(ii) Generalization to mixed state prior:

According to assumption 1 the prior p is invariant under $G_{\mathcal{A}}$ and according to assumption 3', it can be realized by first sampling a pure state from $\tilde{p}$ and then applying channel $\mathcal{N}$ to the state where $\mathcal{N}$ is both $G_{\mathcal{A}}$ covariant and noiseless on $\mathcal{A}$. Then one can easily see that the prior $\tilde{p}$ can always be chosen to be invariant under $G_{\mathcal{A}}$. In other words, for any given prior $\tilde{p}$ which satisfies the above properties there exists a prior p' defined as

$$p'(\cdot) \equiv \int_{G_{\mathcal{A}}} d\mu(V)\ \tilde{p}(V\cdot V^\dagger) \tag{5.19}$$

which also satisfies these properties, i.e. p' is nonzero only on pure states and furthermore one can realize the prior p by sampling a pure state from p' and then applying the quantum channel $\mathcal{N}$ to the state. In addition to these properties, definition 5.19 guarantees that p' is also invariant under $G_{\mathcal{A}}$.

Now consider the estimation problem which is specified by the parameters $\vec{\mathfrak{s}}$, the prior p and the channel $\mathcal{E}$ which satisfy all the assumptions of theorem 13. We call this *estimation problem (a)*. Now define *estimation problem (b)* via the following modifications of problem (a):

1. We change the prior p to p' defined in Eq. (5.19).

2. We change the parameters $\vec{\mathfrak{s}}$ to $\vec{\mathfrak{s}}'$ where

$$\vec{\mathfrak{s}}'(\cdot) \equiv \vec{\mathfrak{s}}\left(\mathcal{N}(\cdot)\right) \tag{5.20}$$

   and so naturally replace the random variables $\vec{S}$ induced by parameters $\vec{\mathfrak{s}}$ to the random variables $\vec{S}'$ induced by parameters $\vec{\mathfrak{s}}'$.

3. We change the channel $\mathcal{E}$ in the problem (a) to the channel

$$\mathcal{E}' \equiv \mathcal{E} \circ \mathcal{N}^{\otimes n}. \tag{5.21}$$

For any POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ let

$$q^{(a)}_M\left(B|\vec{S}\in\vec{\Delta}\right)$$

be the conditional that in problem (a) an event $B \in \sigma(\Omega)$ happens given $\vec{S} \in \vec{\Delta}$ and similarly

$$q^{(b)}_M\left(B|\vec{S}'\in\vec{\Delta}\right)$$

be the conditional that in problem (b) an event $B \in \sigma(\Omega)$ happens given $\vec{S}' \in \vec{\Delta}$.

Now one can easily see that by the manner in which they are defined, the parameters $\mathfrak{s}'$, prior p' and channel $\mathcal{E}'$ of problem (b) satisfy all the assumptions of the theorem.

On the other hand, since p' is nonzero only for pure states then in the case of problem (b) we can use the result of part (i) of this proof, Eq. (5.18), which implies that for any POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$

$$q_M^{(b)}\left(B|\vec{S}' \in \vec{\Delta}\right) = q^{(b)}_{\mathcal{L}_+\circ\mathcal{E}'^\dagger(M)}\left(B|\vec{S}' \in \vec{\Delta}\right) \qquad (5.22)$$

for all $B \in \sigma(\Omega)$ and for all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability.

Then it can be shown that for any POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$ it holds that

$$q_M^{(a)}\left(B|\vec{S} \in \vec{\Delta}\right) = q_M^{(b)}\left(B|\vec{S}' \in \vec{\Delta}\right) \qquad (5.23)$$

for all $B \in \sigma(\Omega)$ and for all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability. We present the proof of this equality at the end. Now this equality allows us to transform the conditionals for problem (a) to the conditionals for the problem (b). Applying Eq. (5.23) to both sides of Eq.(5.22), we get

$$q_M^{(a)}\left(B|\vec{S} \in \vec{\Delta}\right) = q^{(a)}_{\mathcal{L}_+\circ\mathcal{E}'^\dagger(M)}\left(B|\vec{S} \in \vec{\Delta}\right) \qquad (5.24)$$

Recall that the problem (a) is the original problem in the statement of theorem. So, defining

$$M' \equiv \mathcal{L}_+ \circ \mathcal{E}'^\dagger(M) = \mathcal{L}_+ \circ \mathcal{N}^{\dagger\otimes n} \circ \mathcal{E}^\dagger(M)$$

we conclude that in the original problem for arbitrary POVM M, for arbitrary $B \in \sigma(\Omega)$ and for arbitrary $\vec{\Delta}$, it holds that

$$q_M\left(B|\vec{S} \in \vec{\Delta}\right) = q_{M'}\left(B|\vec{S} \in \vec{\Delta}\right) \qquad (5.25)$$

where for all $B \in \sigma(\Omega)$, $M'(B)$ is in $\mathcal{A}^{\otimes n}$ as it is claimed in the theorem.

So it remains only to prove that Eq.(5.23) holds. Let $\vec{\Delta} \subseteq \mathbb{R}^l$, and define probability measures

$$\Pr^{(a)}\left(\vec{S} \in \vec{\Delta}\right) \equiv \int_{\vec{\mathfrak{s}}(\rho)\in\vec{\Delta}} d\rho\ p(\rho) \qquad \text{and,}$$
$$\Pr^{(b)}\left(\vec{S}' \in \vec{\Delta}\right) \equiv \int_{\vec{\mathfrak{s}}'(\rho)\in\vec{\Delta}} d\rho\ p'(\rho).$$

Note that

$$q_M^{(a)}(B|\vec{S} \in \vec{\Delta}) \equiv \frac{\int_{\vec{S}\in\vec{\Delta}} d\rho p(\rho)\ q_M^{(a)}(B|\rho)}{\Pr^{(a)}\left(\vec{S} \in \vec{\Delta}\right)} \qquad \text{and,}$$

$$q_M^{(b)}(B|\vec{S}' \in \vec{\Delta}) \equiv \frac{\int_{\vec{S}'\in\vec{\Delta}} d\rho p'(\rho)\ q_M^{(b)}(B|\rho)}{\Pr^{(b)}\left(\vec{S}' \in \vec{\Delta}\right)}$$

Now using the definition $\mathfrak{s}'(\cdot) \equiv \mathfrak{s}(\mathcal{N}(\cdot))$ from Eq. (5.20), we get

$$\Pr^{(b)}\left(\vec{S}' \in \vec{\Delta}\right) = \int_{\vec{\mathfrak{s}}(\mathcal{N}(\rho))\in\vec{\Delta}} d\rho\ p'(\rho)$$
$$= \int_{\vec{\mathfrak{s}}(\sigma)\in\vec{\Delta}} d\sigma \int_{\mathcal{N}(\rho)=\sigma} d\rho\ p'(\rho),$$

where we have decomposed the integral over $\vec{\mathfrak{s}}(\mathcal{N}(\rho)) \in \vec{\Delta}$ to an integral over all ρ for which $\mathcal{N}(\rho) = \sigma$ and an integral over all σ for which $\vec{\mathfrak{s}}(\sigma) \in \vec{\Delta}$. Finally, since sampling a pure state from p' and applying the channel $\mathcal{N}$ to it realizes the prior p, we conclude that

$$\Pr^{(b)}\left(\vec{S}' \in \vec{\Delta}\right) = \int_{\vec{\mathfrak{s}}(\sigma)\in\vec{\Delta}} d\sigma\ p(\sigma). \qquad (5.26)$$

But this is the definition of $\Pr^{(a)}\left(\vec{S} \in \vec{\Delta}\right)$ and so we conclude that

$$\Pr^{(b)}\left(\vec{S}' \in \vec{\Delta}\right) = \Pr^{(a)}\left(\vec{S} \in \vec{\Delta}\right), \qquad (5.27)$$

Using exactly the same technique for

$$q_M^{(b)}(B|\rho) = \mathrm{tr}\left(\mathcal{E}'(\rho^{\otimes n})M(B)\right) \qquad \text{and}$$
$$q_M^{(a)}(B|\rho) = \mathrm{tr}\left(\mathcal{E}(\rho^{\otimes n})M(B)\right)$$

and the definition $\mathcal{E}' \equiv \mathcal{E}\circ\mathcal{N}^{\otimes n}$, Eq. (5.21), we can prove that

$$\int_{\vec{S}\in\vec{\Delta}} d\rho\ p(\rho)\ q_M^{(a)}(B|\rho) = \int_{\vec{S}'\in\vec{\Delta}} d\rho\ p'(\rho)\ q_M^{(b)}(B|\rho) \qquad (5.28)$$

Eqs. (5.28) and (5.27) together imply Eq.(5.23). This completes the proof. ■

VI. SINGLE-COPY ESTIMATION PROBLEMS FOR BIPARTITE SYSTEMS

Previously in this article, the distinction between global and local symmetries was relative to the partitioning of the total system into n copies of the system of interest. However, one can also consider estimation problems where the estimator gets only a single copy of the system of interest, and the distinction between global and local symmetries is relative to the partitioning of the system of interest into its components. This case can be significantly different because the components of the system of interest need not correspond to copies of a single state. Indeed, they could even be entangled.

In particular, we consider the case where the system has only *two* components. This case allows us to obtain particularly strong constraints on the optimal measurement because the permutation group on two systems has only irreducible representations over the symmetric and antisymmetric subspaces and our duality only permits an inference from global symmetry to local symmetry within the symmetric and antisymmetric subspaces (as shown by the counterexample from Appendix C).

We begin with some notation. The canonical representation of the permutation group on the pair is $\mathbf{P}(\mathcal{S}_2) \equiv \{I, \mathrm{Swap}\}$, where I is the identity operator on $(\mathbb{C}^d)^{\otimes 2}$ and Swap is the unitary which exchanges the state of the

two systems, i.e. $\mathrm{Swap}(|\psi\rangle|\phi\rangle) = |\phi\rangle|\psi\rangle$. Under $\mathbf{P}(\mathcal{S}_2)$, the space $(\mathbb{C}^d)^{\otimes 2}$ decomposes as

$$(\mathbb{C}^d)^{\otimes 2} \cong [(\mathbb{C}^d)^{\otimes 2}]_+ \oplus [(\mathbb{C}^d)^{\otimes 2}]_- \tag{6.1}$$

Also, for any subgroup $H \subseteq \mathrm{U}(d)$, the collective representation of H on the pair of systems is denoted by $\mathbf{Q}(G_{\mathcal{A}}) \equiv \{V^{\otimes 2} : V \in G_{\mathcal{A}}\}$.

We are now in a position to state our result.

Scenario: Suppose that Alice randomly chooses an unknown state ρ from the density operators in $\mathrm{End}\left((\mathbb{C}^d)^{\otimes 2}\right)$ according to some probability density p and sends a single system in the state ρ to Bob. Let $\vec{\mathfrak{s}}(\cdot) = \left(\mathfrak{s}^{(1)}(\cdot), \cdots, \mathfrak{s}^{(l)}(\cdot)\right)$ be an arbitrary set of functions where $\mathfrak{s}^{(i)} : \mathrm{supp}(p) \to \mathbb{R}$, and let $\vec{S}$ be the random variables defined as $\vec{S} \equiv \vec{\mathfrak{s}}(\rho)$ where ρ is the random state Alice chooses. We refer to $\vec{\mathfrak{s}}$ as the *parameters*.

Recalling our earlier definitions, Eqs. (5.3) and (5.4), of what it means for a prior p and a vector of parameters $\vec{\mathfrak{s}}$ to have a symmetry, we can state our result as follows:

Theorem 17 *Let $\mathcal{A} \subseteq End(\mathbb{C}^d)$ be a von Neumann algebra with the gauge group $G_{\mathcal{A}}$. Then assuming that the prior p and the vector of parameters $\vec{\mathfrak{s}}$*

1. *have $\boldsymbol{Q}(G_{\mathcal{A}})$ as a symmetry,*
2. *have $\boldsymbol{P}(\mathcal{S}_2)$ as a symmetry,*

then for any given measurement with POVM $M : \sigma(\Omega) \to End((\mathbb{C}^d)^{\otimes 2})$, there is another measurement whose POVM is of the form

$$M' \equiv \Pi_+ M_+ \Pi_+ + \Pi_- M_- \Pi_-$$

where $M_\pm : \sigma(\Omega) \to \mathcal{A}^{\otimes 2}$ are POVMs, such that M' is as informative about $\vec{S}$ as M is, i.e.,

$$q_M\left(B|\vec{S} \in \vec{\Delta}\right) = q_{M'}\left(B|\vec{S} \in \vec{\Delta}\right) \tag{6.2}$$

for all $B \in \sigma(\Omega)$ and all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability.

The proof is provided at the end of this section. Note that unlike theorems 11 and 13, the prior is not presumed to have support only on the pure states nor to be a gauge distortion of one that does.

Remark 18 *The measurement M' described in the above theorem can be implemented as follows: first perform the measurement which projects to the symmetric/anti-symmetric subspace (the projective measurement described by projectors $\{\Pi_+, \Pi_-\}$) and then, depending on the outcome of this measurement, perform either measurement M_+ or M_- where both have local symmetry with respect to $G_{\mathcal{A}}$. The outcome of measurement M' is the outcome of whichever of these measurements was performed.*

A. Example

Suppose that the prior over the pair of systems has support only on product states $\rho_1 \otimes \rho_2$ where $\rho_1, \rho_2 \in \mathrm{End}(\mathbb{C}^d)$ and that it corresponds to choosing ρ_1 and ρ_2 independently according to a prior p_0, so that the joint prior has the form $p(\rho_1 \otimes \rho_2) = p_0(\rho_1)p_0(\rho_2)$. Assume further that $p_0(\rho)$ only depends on the eigenvalues of ρ, so that $p_0(\cdot) = p_0\left(V(\cdot)V^\dagger\right)$ for arbitrary $V \in \mathrm{U}(d)$, i.e., p_0 has $\mathrm{U}(d)$ as a symmetry. It follows that the prior p on the pair has $\mathbf{Q}(\mathrm{U}(d))$ as a symmetry, and consequently it also has $\mathbf{Q}(H)$ as a symmetry for any subgroup H of $\mathrm{U}(d)$. Moreover, the prior p is invariant under permutations, i.e. it has $\mathbf{P}(\mathcal{S}_2)$ as a symmetry.

The goal is to estimate the parameter $\mathfrak{s}(\rho_1 \otimes \rho_2) = |\mathrm{tr}\,(A\rho_1) - \mathrm{tr}\,(A\rho_2)\,|$ for some observable A. Let $\mathcal{A}$ denote the algebra generated by $\{I, A\}$ and let $G_{\mathcal{A}}$ denote the associated gauge group. It is clear that $\mathfrak{s}$ has $\mathbf{Q}(G_{\mathcal{A}})$ as a symmetry. Furthermore, $\mathfrak{s}$ is invariant under a swap of the pair of systems and therefore has $\mathbf{P}(\mathcal{S}_2)$ as a symmetry as well. The parameter $\mathfrak{s}$ therefore satisfies the assumptions of the above theorem for the gauge group $G_{\mathcal{A}}$. Furthermore, because $G_{\mathcal{A}}$ is a subgroup of $\mathrm{U}(d)$, the prior p satisfies the assumptions of the above theorem as well.

So, for any figure of merit that can be defined as a functional acting on $q_M\left(B|S = s_0\right)$, the optimal estimation strategy corresponds to a POVM M' of the form described in the theorem. In our example, such a measurement has a particularly simple form. First, note that because the two POVMs M_+ and M_- have local symmetry with respect to $G_{\mathcal{A}}$ and because $\mathcal{A}$ is commutative, using proposition 9, we can conclude that M_+ and M_- can both be realized by measuring a Hermitian generator of $\mathcal{A}$ (e.g. the operator A) individually on each system and performing a classical processing of the outcome. This means that in the case of this example, the POVM M' described in the theorem can be realized by (i) performing the measurement which projects the state into the symmetric and antisymmetric subspaces, (ii) measuring the observable A individually on each system and (iii) generating the outcome by a classical processing of the outcomes of these measurements. So for all such M's, the measurements are fixed and the part which is different is just the classical processing.

The same result holds for any other parameter which is invariant with respect to the exchange of the pair of systems and can be expressed in terms of an operator A, such as $\mathfrak{s}(\rho_1 \otimes \rho_2) = \mathrm{tr}\,(A\rho_1) + \mathrm{tr}\,(A\rho_2)$ or more complicated parameters such as $\mathfrak{s}(\rho_1 \otimes \rho_2) = \mathrm{tr}\left(A\rho_1^k\rho_2^k\right) + \mathrm{tr}\left(A\rho_2^k\rho_1^k\right)$ for some integer k.

B. Proof of theorem 17

Proof. (Theorem 17) We need to apply lemma 16 in its special case where $n = 1$ and the Hilbert space of a single copy (which was denoted by $\mathbb{C}^d$ in the statement

of lemma) is $\mathbb{C}^d \otimes \mathbb{C}^d$. The symmetry of the problem, denoted by $H \subseteq \mathrm{U}(d^2)$, is the group generated by $\mathbf{Q}(G_{\mathcal{A}})$ and $\mathbf{P}(\mathcal{S}_2)$ together. Then lemma 16 implies that for any POVM $M_{\pm} : \sigma(\Omega) \to \mathrm{End}(\mathbb{C}^d \otimes \mathbb{C}^d)$ there is a POVM

$$\tilde{M} \equiv \mathcal{T}_H(M) = \int_H d\mu(V)\ VMV^{\dagger}$$

such that

$$q_{\tilde{M}}\left(B|\vec{S} \in \vec{\Delta}\right) = q_M\left(B|\vec{S} \in \vec{\Delta}\right)$$

for all $B \in \sigma(\Omega)$ and all l-dimensional intervals $\vec{\Delta}$ which are assigned nonzero probability. Now the above definition implies that $\tilde{M}$ is invariant under permutation i.e. $\tilde{M} = \mathrm{Swap}[\tilde{M}]\mathrm{Swap}$. This implies that[1]

$$\tilde{M} = \Pi_+\tilde{M}\Pi_+ + \Pi_-\tilde{M}\Pi_-.$$

$\tilde{M}$ also has global symmetry with respect to the gauge group $G_{\mathcal{A}}$, i.e. it commutes with $\mathbf{Q}(G_{\mathcal{A}})$. Now corollary 10 implies that for states whose supports are restricted to the symmetric/anti-symmetric subspaces a measurement with global symmetry with respect to gauge group G_A can be simulated by a measurement whose POVM has local symmetry (and so its POVM elements are in $\mathcal{A}\otimes\mathcal{A}$). Therefore there exists POVMs M_+ and M_- where $M_{\pm} : \sigma(\Omega) \to \mathcal{A} \otimes \mathcal{A}$ such that

$$\Pi_+M_+\Pi_+ = \Pi_+\tilde{M}\Pi_+ \quad \text{and} \quad \Pi_-M_-\Pi_- = \Pi_-\tilde{M}\Pi_-$$

An example of $M_{\pm}$ is $\mathcal{L}_{\pm}(\tilde{M})$. Also since $\Pi_{\pm}\mathcal{L}_{\pm}(\tilde{M})\Pi_{\pm} = \Pi_{\pm}\mathcal{L}_{\pm}(M)\Pi_{\pm}$, it follows that $\mathcal{L}_{\pm}(M)$ is also an example of $M_{\pm}$. This completes the proof. ■

Acknowledgments

We acknowledge helpful discussions with Giulio Chiribella. Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. I. M. is supported by a Mike and Ophelia Lazaridis fellowship and NSERC.

[1] Therefore one can precede the measurement of $\tilde{M}$ with the measurement which projects to the symmetric/anti-symmetric subspace (the projective measurement described by projectors $\{\Pi_+, \Pi_-\}$) without changing its statistics. If we like, we can associate the projection onto symmetric/antisymmetric subspaces with the state preparation rather than with the measurement, in which case, the states can be considered to be restricted to the symmetric/antisymmetric subspaces.

Appendix A: Proofs of lemma 5 and theorem 8

Throughout these proofs we use the superoperator $\mathcal{T}_{\mathcal{S}_n} : \mathrm{End}((\mathbb{C}^d)^{\otimes n}) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$

$$\mathcal{T}_{\mathcal{S}_n}(\cdot) \equiv \frac{1}{n!}\sum_{s\in\mathcal{S}_n} \mathbf{P}(s)(\cdot)\mathbf{P}^{\dagger}(s) \tag{A1}$$

which maps any operator in $\mathrm{End}((\mathbb{C}^d)^{\otimes n})$ to its symmetrized version (under permutation).

Proof. (**lemma 5**) First note that $\mathrm{Alg}\{\mathbf{Q}(G)\} \subseteq \mathrm{Alg}\{G^{\times n}\}$ and furthermore all elements of $\mathrm{Alg}\{\mathbf{Q}(G)\}$ are permutationally invariant. So $\mathrm{Alg}\{\mathbf{Q}(G)\}$ is included in the permutationally invariant subalgebra of $\mathrm{Alg}\{G^{\times n}\}$. In the following, we prove the converse inclusion.

We prove this by induction. First we prove that for arbitrary $V_0 \in G$, the subspace spanned by $\mathcal{T}_{\mathcal{S}_n}(V_0 \otimes I^{\otimes(n-1)})$ is in $\mathrm{Alg}\{\mathbf{Q}(G)\}$. Then by induction we prove it is true for $\mathcal{T}_{\mathcal{S}_n}(V_0 \otimes \cdots \otimes V_n)$ for arbitrary $V_i \in G : i = 1 \cdots n$ which proves the claim.

For arbitrary unitary $V_0 \in G$, clearly $V_0 + V_0^{\dagger}$ and $i(V_0 - V_0^{\dagger})$ are both Hermitian operators which commute with G' (the centralizer of G). Therefore, all operators of the form $V_0(\theta,\phi) \equiv \exp\left[i\theta(V_0 + V_0^{\dagger}) + \phi(V_0 - V_0^{\dagger})\right]$, for arbitrary real numbers θ and ϕ are unitary and commute with G'. By virtue of being a gauge group, G includes all unitaries which commute with G', and it therefore follows that $V_0(\theta,\phi) \in G$. We can easily see that

$$\frac{1}{2}\left(\frac{\partial}{\partial\phi} - i\frac{\partial}{\partial\theta}\right)_{|\theta=\phi=0} V_0(\theta,\phi) = V_0 \tag{A2}$$

This implies that

$$\frac{1}{2}\left(\frac{\partial}{\partial\phi} - i\frac{\partial}{\partial\theta}\right)_{|\theta=\phi=0} V_0^{\otimes n}(\theta,\phi) = \sum_k V_0^{(k)} \tag{A3}$$

where $V_0^{(k)} \equiv I^{\otimes(k-1)} \otimes V_0 \otimes I^{\otimes(n-k)}$. This means that for arbitrary $V_0 \in G$ the operator $\mathcal{T}_{\mathcal{S}_n}(V_0 \otimes I^{\otimes(n-1)})$ is in $\mathrm{Alg}\{\mathbf{Q}(G)\}$.

Next we assume that

$$\mathcal{T}_{\mathcal{S}_n}(V_0 \otimes \cdots \otimes V_{k-1} \otimes I^{\otimes(n-k)})$$

is in $\mathrm{Alg}\{\mathbf{Q}(G)\}$ for arbitrary $V_i \in G : i = 0 \cdots k-1$. This implies that for arbitrary $V_k \in G$

$$\mathcal{T}_{\mathcal{S}_n}(V_0 \otimes \cdots V_{k-1} \otimes I^{\otimes(n-k)})\mathcal{T}_{\mathcal{S}_n}(V_k \otimes I^{\otimes(n-1)})$$

is in $\mathrm{Alg}\{\mathbf{Q}(G)\}$. Expanding this, one can easily see that it can be written as

$$c_1\mathcal{T}_{\mathcal{S}_n}(V_0 \otimes \cdots V_{k-1} \otimes V_k \otimes I^{\otimes(n-k-1)}) + c_2\mathcal{T}_{\mathcal{S}_n}(U_0 \otimes \cdots U_{k-1} \otimes I^{\otimes(n-k)})$$

for some nonzero coefficients c_1, c_2 and unitaries $U_i \in G : i = 0 \cdots k-1$. Now since the sum and the second term

each are in the span of $\mathrm{Alg}\{\mathbf{Q}(G)\}$ then we conclude that the first term is also in $\mathrm{Alg}\{\mathbf{Q}(G)\}$. Note that k and $V_i \in G : i = 0 \cdots k$ are arbitrary. So by induction we have the lemma. ■

Proof. (theorem 8)

For any arbitrary operator $V \in \mathrm{U}(d)$, $V^{\otimes n}$ commutes with $\Pi_\pm$. So the condition $\forall V \in G_{\mathcal{A}} : \Pi_\pm M \Pi_\pm \mathbf{Q}(V) = \mathbf{Q}(V)\Pi_\pm M \Pi_\pm$ is equivalent to

$$\Pi_\pm M \Pi_\pm \mathbf{Q}(V) \Pi_\pm = \Pi_\pm \mathbf{Q}(V) \Pi_\pm M \Pi_\pm \tag{A4}$$

This holds for arbitrary $V \in G_{\mathcal{A}}$. So it implies that for any operator X in $\mathrm{Alg}\{\mathbf{Q}(G_{\mathcal{A}})\}$ we have

$$\Pi_\pm M \Pi_\pm X \Pi_\pm = \Pi_\pm X \Pi_\pm M \Pi_\pm \tag{A5}$$

According to lemma 5, $\mathrm{Alg}\{\mathbf{Q}(G_{\mathcal{A}})\}$ is equal to the span of the permutationally invariant subspace of $G_{\mathcal{A}}^{\times n}$. Consider $V_1 \otimes \cdots \otimes V_n$ an arbitrary element of $G_{\mathcal{A}}^{\times n}$. Since $\mathcal{T}_{\mathcal{S}_n}(V_1 \otimes \cdots \otimes V_n)$ is in the permutationally invariant subspace of the span of $G_{\mathcal{A}}^{\times n}$, it satisfies Eq.(A5) and so

$$\begin{aligned} &\Pi_\pm M \Pi_\pm \left[\mathcal{T}_{\mathcal{S}_n}(V_1 \otimes \cdots \otimes V_n)\right] \Pi_\pm = \\ &\qquad \Pi_\pm \left[\mathcal{T}_{\mathcal{S}_n}(V_1 \otimes \cdots \otimes V_n)\right] \Pi_\pm M \Pi_\pm \end{aligned} \tag{A6}$$

For arbitrary permutation $s \in \mathcal{S}_n$, $\mathbf{P}(s)\Pi_\pm = \Pi_\pm \mathbf{P}(s) = \eta \Pi_\pm$ for some $\eta \in \{\pm 1\}$. Therefore Eq.(A6) implies

$$\begin{aligned} &\Pi_\pm M \Pi_\pm \left[V_1 \otimes \cdots \otimes V_n\right] \Pi_\pm \\ &\qquad = \Pi_\pm \left[V_1 \otimes \cdots \otimes V_n\right] \Pi_\pm M \Pi_\pm \end{aligned}$$

We multiply by $[V_1^\dagger \otimes \cdots \otimes V_n^\dagger]\Pi_\pm$ on the right on both sides of the above equality to obtain

$$\begin{aligned} &\Pi_\pm M \Pi_\pm \left[(V_1 \otimes \cdots \otimes V_n)\Pi_\pm (V_1^\dagger \otimes \cdots \otimes V_n^\dagger)\right] \Pi_\pm \\ &\quad = \Pi_\pm \left[(V_1 \otimes \cdots \otimes V_n)\Pi_\pm M \Pi_\pm (V_1^\dagger \otimes \cdots \otimes V_n^\dagger)\right] \Pi_\pm \end{aligned}$$

Now suppose on both sides we integrate over all elements of $G_{\mathcal{A}}^{\times n}$ using the Haar measure. Then the above equality implies

$$\Pi_\pm M \Pi_\pm [\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)] \Pi_\pm = \Pi_\pm \mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm) \Pi_\pm \tag{A7}$$

Now we demonstrate how one can write $\Pi_\pm M \Pi_\pm$ as $\Pi_\pm \mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm)\Pi_\pm$ times the inverse of $\Pi_\pm [\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)]\Pi_\pm$.

Consider $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)$ and $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm)$ on the left and right hand sides of the above equality. First of all, since $\Pi_\pm$ and $\Pi_\pm M \Pi_\pm$ are both permutationally invariant then both $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)$ and $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm)$ are permutationally invariant. Furthermore, since these two operators also commute with $G^{\times n}$ then corollary 6 implies that they are both in $\mathrm{Alg}\{\mathbf{Q}(G'_{\mathcal{A}})\}$. Second, since $\Pi_\pm$ commutes with $\mathbf{Q}(G'_{\mathcal{A}})$ in the case of $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)$ we have another symmetry: $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)$ commutes with $\mathbf{Q}(G'_{\mathcal{A}})$. Considering this fact together with the fact that $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)$ is in $\mathrm{Alg}\{\mathbf{Q}(G'_{\mathcal{A}})\}$ we conclude that $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm)$ should have the following form

$$\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm) = \bigoplus_\mu p_\mu P_\mu \tag{A8}$$

where μ labels all the irreps of $G'_{\mathcal{A}}$ which shows up in the representation $\mathbf{Q}(G'_{\mathcal{A}})$ and P_μ is the projector to these irreps and by virtue of $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}$ being a completely positive map, all p_μ's are non-negative. Let Γ be the set of all irreps of $G'_{\mathcal{A}}$ for which p_μ is nonzero. So we can write Eq.(A7) as

$$\Pi_\pm M \Pi_\pm \left(\bigoplus_{\mu \in \Gamma} p_\mu P_\mu\right) \Pi_\pm = \Pi_\pm \left[\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm)\right] \Pi_\pm \tag{A9}$$

Now consider the inverse of $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm) = \bigoplus_{\mu \in \Gamma}(p_\mu P_\mu)$ on its support, i.e., the operator

$$\bigoplus_{\mu \in \Gamma} p_\mu^{-1} P_\mu$$

By multiplying both sides of Eq.(A9) on the right with this operator and using the facts that

1. $\Pi_\pm$ commutes with $\mathbf{Q}(G'_{\mathcal{A}})$ and so it commutes with all P_μ's,

2. $$\Pi_\pm \left(\bigoplus_{\mu \in \Gamma} P_\mu\right) = \left(\bigoplus_{\mu \in \Gamma} P_\mu\right) \Pi_\pm = \Pi_\pm \tag{A10}$$

   which is true because all P_μ's commute with $\Pi_\pm$ and Eq.(A8) implies that the support of $\Pi_\pm$ is a subspace of the support of $\bigoplus_{\mu \in \Gamma} P_\mu$ and

3. $$\forall \mu : \; P_\mu \mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm) = \mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm) P_\mu$$

   which is true because $\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm)$ is in the span of $\mathbf{Q}(G'_{\mathcal{A}})$

we get

$$\Pi_\pm M \Pi_\pm = \Pi_\pm \left(\bigoplus_{\mu \in \Gamma} p_\mu^{-1} P_\mu \left[\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm M \Pi_\pm)\right] P_\mu\right) \Pi_\pm \tag{A11}$$

Therefore, defining $\Phi_\pm$ as

$$\Phi_\pm(\cdot) \equiv \bigoplus_\mu p_{\mu,\pm}^{-1} \; P_\mu [\mathcal{T}_{G_{\mathcal{A}}}^{\otimes n}(\Pi_\pm(\cdot)\Pi_\pm)] P_\mu \tag{A12}$$

with summation over all μ's for which p_μ is nonzero, we infer that

$$\Pi_\pm M \Pi_\pm = \Pi_\pm \Phi_\pm(M) \Pi_\pm \tag{A13}$$

Because all P_μ's and $\mathcal{T}_{G'_\mathcal{A}}^{\otimes n}(\Pi_\pm(\cdot)\Pi_\pm)$ are in $\mathrm{Alg}\{\mathbf{Q}(G'_\mathcal{A})\}$, the image of $\Phi_\pm$ is as well. Note that since $G'_\mathcal{A} \subset \mathcal{A}$ this means that the image of $\Phi_\pm$ is in the permutationally invariant subalgebra of $\mathcal{A}^{\otimes n}$. Now defining $\mathcal{L}_\pm$ in terms of $\Phi_\pm$ via

$$\mathcal{L}_\pm(\cdot) \equiv \Phi_\pm(\cdot) + [I^{\otimes n} - \Phi_\pm(I^{\otimes n})]\mathrm{tr}(\cdot)/d^n$$

we can infer the same properties for $\mathcal{L}_\pm$. First note that

$$\Phi_\pm(I^{\otimes n}) = \bigoplus_{\mu\in\Gamma} P_\mu$$

which together with Eq.(A10) implies that $\Pi_\pm[I^{\otimes n} - \Phi_\pm(I^{\otimes n})]\Pi_\pm = 0$. This together with Eq.(A13) and definition of $\mathcal{L}_\pm$ implies

$$\Pi_\pm M \Pi_\pm = \Pi_\pm \mathcal{L}_\pm(M)\Pi_\pm\,, \tag{A14}$$

which is the third claim of theorem 8. Furthermore since the image of $\Phi_\pm$ is in the permutationally invariant subalgebra of $\mathcal{A}^{\otimes n}$ and since $\mathcal{A}$, being a von-Neumann algebra, includes identity, it follows that $I^{\otimes n} - \Phi_\pm(I^{\otimes n})$ is in the permutationally invariant subalgebra of $\mathcal{A}^{\otimes n}$. This implies that the image of $\mathcal{L}_\pm(\cdot)$ is in this subalgebra, which is the second claim of theorem 8.

Furthermore, noting that $\mathcal{T}_{G'_\mathcal{A}}^{\otimes n}$ is completely positive and the p_μ^{-1}'s are all positive numbers we can conclude that $\Phi_\pm$ as a combination of completely positive maps is completely positive. This together with the fact that $I^{\otimes n} - \Phi_\pm(I^{\otimes n})$ is a projector (and so a positive operator) implies that $\mathcal{L}_\pm$ is completely positive. Finally, It is straightforward to verify that $\mathcal{L}_\pm(I^{\otimes n}) = I^{\otimes n}$, so that it is unital which proves the first claim of theorem 8. ■

Appendix B: Global symmetry with respect to non-gauge groups

We demonstrate here that a group that does not have the gauge property does not yield a dual reductive pair in the manner specified by theorems 3 and 7. That is, we present an example for a non-gauge group $H \in \mathrm{U}(d)$ for which the commutant of the algebra spanned by $\mathbf{Q}(H)$ in $\mathrm{End}((\mathbb{C}^d)^{\otimes n})$ is larger than the algebra spanned by $\langle (H')^{\times n}, \mathbf{P}(\mathcal{S}_n)\rangle$. (Recall that for any group $H \subseteq \mathrm{U}(d)$ it always holds that $\mathrm{Alg}\{(H')^{\times 2}, \mathbf{P}(\mathcal{S}_2)\} \subseteq \mathrm{Comm}\{\mathbf{Q}(H)\}$).

As a simple example, consider $d = 3$, $n = 2$ where the group H is the $j = 1$ irreducible representation of SU(2) which is a subgroup of U(3). This group is not a gauge group: Schur's lemma implies that $H' = \{e^{i\theta} I\}$ where $\theta \in (0, 2\pi]$ and I is identity on $\mathbb{C}^3$ and so $H'' = \mathrm{U}(3) \neq H$.

So

$$\begin{aligned}\mathrm{Alg}\{(H')^{\times 2}, \mathbf{P}(\mathfrak{S}_2)\} &= \mathrm{Alg}\{\mathbf{P}(\mathfrak{S}_2)\} \\ &= \{c_+\Pi_+ + c_-\Pi_- : c_\pm \in \mathbb{C}\}\end{aligned}$$

where Π_+ and Π_- are respectively the projectors to the symmetric and anti-symmetric subspace of $(\mathbb{C}^3)^{\otimes 2}$. On the other hand, one can easily see that $\mathrm{Comm}\{\mathbf{Q}(H)\}$, the algebra of operators commuting with $\mathbf{Q}(H)$, is

$$\{c_0\Pi_{j=0} + c_1\Pi_{j=1} + c_2\Pi_{j=2}, c_{0,1,2} \in \mathbb{C}\}$$

where Π_j is the projector to the subspace of $(\mathbb{C}^3)^{\otimes 2}$ with total angular momentum j. Therefore the algebra of operators commuting with $\mathbf{Q}(H)$ is larger than $\mathrm{Alg}\{(H')^{\times 2}, \mathbf{P}(\mathcal{S}_2)\}$.

Appendix C: Lack of duality outside the symmetric and antisymmetric subspaces

Here, we show that the restriction to the symmetric and anti-symmetric subspaces plays an essential role in theorem 7 and the other results of section III D. Recall that theorem 7 implies that for symmetric and anti-symmetric subspaces of $(\mathbb{C}^d)^{\otimes n}$, denoted by $\left[(\mathbb{C}^d)^{\otimes n}\right]_\pm$, and for any gauge group $G \subseteq \mathrm{U}(d)$ it holds that

$$\mathrm{Alg}\{\Pi_\pm \mathbf{Q}(G')\Pi_\pm\} = \mathrm{Comm}\{\Pi_\pm \mathbf{Q}(G)\Pi_\pm\}$$

Clearly for any other irrep λ of $\mathcal{S}_n$ it holds that

$$\mathrm{Alg}\{\Pi_\lambda \mathbf{Q}(G')\Pi_\lambda\} \subseteq \mathrm{Comm}\{\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda\}$$

where by $\mathrm{Comm}\{\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda\}$ we mean the commutant of $\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda \in \mathrm{End}(\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda)$ in $\mathrm{End}(\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda)$. However, for arbitrary λ, $\mathrm{Comm}\{\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda\}$ is not necessarily invariant under permutation, while $\mathrm{Alg}\{\Pi_\lambda \mathbf{Q}(G')\Pi_\lambda\}$ is always permutationally invariant. So, as we discussed in section III D a natural generalization of theorem 7 will be

$$\begin{aligned}&\mathrm{Alg}\{\Pi_\lambda \mathbf{Q}(G')\Pi_\lambda\} = \\ &\qquad \mathrm{Comm}\{\Pi_\pm \mathbf{Q}(G)\Pi_\pm\} \cap \mathrm{Comm}\{\Pi_\lambda \mathbf{P}(\mathcal{S}_n)\Pi_\lambda\}\end{aligned}$$

where Π_λ is the projector to the subspace of $(\mathbb{C}^d)^{\otimes n}$ which carries irrep λ of $\mathcal{S}_n$. From theorem 7 we know that for the special case of 1-d representations of $\mathcal{S}_n$, i.e. for symmetric and anti-symmetric subspaces, the above equality hold. Here, we build an explicit counter-example to this equality for other irreps of $\mathcal{S}_n$.

First, notice that the action of $\mathbf{Q}(G)$, $\mathbf{Q}(G')$ and $\mathbf{P}(\mathcal{S}_n)$ on $(\mathbb{C}^d)^{\otimes n}$ all commute with each other. This implies that for irrep λ of $\mathcal{S}_n$ the subspace $\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda$ can be decomposed as

$$\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda \cong \left(\bigoplus_{\mu,\nu} \mathcal{M}_\mu \otimes \mathcal{N}_\nu \otimes \mathbb{C}^{m_{\mu,\nu}}\right) \otimes \mathcal{K}_\lambda \tag{C1}$$

where μ labels irreps of G and ν labels irreps of G' and $\mathbf{Q}(G)$, $\mathbf{Q}(G')$ and $\mathbf{P}(\mathcal{S}_n)$ act nontrivially only on $\mathcal{M}_\mu$, $\mathcal{N}_\nu$ and $\mathcal{K}_\lambda$ respectively. Note that any permutationally invariant operator should be proportional to identity on the subsystem $\mathcal{K}_\lambda$.

Now to build the counterexample we find two gauge groups G and G' for which there is no one-to-one relation between the irreps of G and G' which show up in $\left[(\mathbb{C}^d)^{\otimes n}\right]_\lambda$. In other words, we find an example in which there is some irrep μ of G for which $m_{\mu,\nu}$ is nonzero for more than one ν (irrep of G'). This in turn will imply that there exist permutationally invariant operators $\Pi_\lambda M \Pi_\lambda$ which commute with $\Pi_\lambda \mathbf{Q}(G)\Pi_\lambda$ but are not block diagonal between $\mathcal{N}_{\nu_1}$ and $\mathcal{N}_{\nu_2}$ for two different irrep ν_1 and ν_2 of G'. This implies that $\Pi_\lambda M \Pi_\lambda$ cannot be in $\mathrm{Alg}\{\Pi_\lambda \mathbf{Q}(G')\Pi_\lambda\}$.

Note that from theorem 7 we know that in the specific case where irrep λ is a 1-d representations of $\mathcal{S}_n$ the conjecture holds. So to find a counter-example we need to look at $n > 2$ where the permutation group can have representations other than the symmetric and anti-symmetric. In the following, we present a counter-example in the case of $n = 3$. In this case the permutation group $\mathcal{S}_3$ has a two dimensional irrep denoted by λ_2.

1. Counter-example

Consider the Hilbert space $\mathbb{C}^4 \cong \mathcal{H}_L \otimes \mathcal{H}_R$ where $\mathcal{H}_L$ and $\mathcal{H}_R$ are both isomorphic to $\mathbb{C}^2$. Suppose $G = \{V \otimes I : V \in U(2)\}$, i.e. the group of all unitaries which act trivially on $\mathcal{H}_R$. Clearly G' is the set of all unitaries acting trivially on $\mathcal{H}_L$ and so $G = (G')'$. Note that both G and G' are isomorphic to U(2). So rather than talking about irreps of G and G' we will talk about irreps of U(2).

Using decomposition $(\mathbb{C}^4)^{\otimes 3} \cong \mathcal{H}_L^{\otimes 3} \otimes \mathcal{H}_R^{\otimes 3}$ we can think of the collective representation of G and G' on $(\mathbb{C}^4)^{\otimes 3}$ as

$$\mathbf{Q}(V \otimes I) = \mathbf{Q}_L(V) \otimes I_R$$

and

$$\mathbf{Q}(I \otimes V) = I_L \otimes \mathbf{Q}_R(V)$$

respectively where $\mathbf{Q}_{L/R}(\mathrm{U}(2))$ can be thought as the collective representation of U(2) on $\mathcal{H}_{L/R}^{\otimes 3}$ and $I_{L/R}$ is the identity operator on $\mathcal{H}_{L/R}^{\otimes 3}$.

Similarly we can think of the canonical representation of $\mathcal{S}_3$ on $(\mathbb{C}^4)^{\otimes 3}$ as

$$\mathbf{P}(s \in \mathcal{S}_3) = \mathbf{P}_L(s) \otimes \mathbf{P}_R(s)$$

where $\mathbf{P}_L(\mathcal{S}_3)$ and $\mathbf{P}_R(\mathcal{S}_3)$ are the canonical representation of $\mathcal{S}_3$ on $\mathcal{H}_L^{\otimes 3}$ and $\mathcal{H}_R^{\otimes 3}$ respectively.

Now according to Schur-Weyl duality there is a one to one relation between the irreps of U(2) which show up in representation $\mathbf{Q}_{L/R}(\mathrm{U}(2))$ on $(\mathcal{H}_{L/R})^{\otimes 3}$ and irreps of $\mathcal{S}_3$ which show up in representation $\mathbf{P}_{L/R}(\mathcal{S}_3)$ on $(\mathcal{H}_{L/R})^{\otimes 3}$. Note that under the action of $\mathcal{S}_3$, $\mathcal{H}_{L/R}^{\otimes 3}$ decomposes as

$$\mathcal{H}_{L/R}^{\otimes 3} \cong \left[\mathcal{H}_{L/R}^{\otimes 3}\right]_+ \oplus \left[\mathcal{H}_{L/R}^{\otimes 3}\right]_{\lambda_2}$$

(the anti-symmetric irrep of $\mathcal{S}_3$ does not exist in this representation.) Now Schur-Weyl duality implies that in the representation $\mathbf{Q}_{L/R}(\mathrm{U}(2))$ of U(2) only one irrep of U(2) shows up in the subspace $\left[\mathcal{H}_{L/R}^{\otimes 3}\right]_+$ and a different one will show up in $\left[\mathcal{H}_{L/R}^{\otimes 3}\right]_{\lambda_2}$.

This implies that there is a one-to-one relation between irreps of U(2) which show up in representation $\mathbf{Q}_L(\mathrm{U}(2)) \otimes I_R$ in the total Hilbert space $(\mathbb{C}^4)^{\otimes 3}$ and irreps of $\mathcal{S}_3$ which show up in the representation $\mathbf{P}_L(\mathcal{S}_3) \otimes I_R$ in the total Hilbert space $(\mathbb{C}^4)^{\otimes 3}$ (though $(\mathbf{P}_L(\mathcal{S}_3) \otimes I_R) \times (\mathbf{Q}_L(\mathrm{U}(2)) \otimes I_R)$ is no longer multiplicity-free). In other words, in representation $\mathbf{Q}_L(\mathrm{U}(2)) \otimes I_R$ only one irrep of U(2) shows up in the subspace

$$\left[\mathcal{H}_L^{\otimes 3}\right]_+ \otimes \left(\left[\mathcal{H}_R^{\otimes 3}\right]_+ \oplus \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}\right)$$

and a different one shows up in

$$\left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2} \otimes \left(\left[\mathcal{H}_R^{\otimes 3}\right]_+ \oplus \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}\right)$$

Similarly, under $I_L \otimes \mathbf{Q}_R(\mathrm{U}(2))$ only one irrep of U(2) shows up in the subspace

$$\left(\left[\mathcal{H}_L^{\otimes 3}\right]_+ \oplus \left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2}\right) \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_+$$

and a different one shows up in

$$\left(\left[\mathcal{H}_L^{\otimes 3}\right]_+ \oplus \left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2}\right) \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}$$

Now we find which parts of these subspaces of $(\mathbb{C}^4)^{\otimes 3}$ form $\left[(\mathbb{C}^4)^{\otimes 3}\right]_{\lambda_2}$ and we show that in this subspace there is no one-to-one relation between irreps of U(2) which show up in the representation $\mathbf{Q}_L(\mathrm{U}(2)) \otimes I_R$ and irreps of U(2) which show up in the representation $I_L \otimes \mathbf{Q}_R(\mathrm{U}(2))$.

To see this consider the total Hilbert space

$$\begin{aligned}(\mathbb{C}^4)^{\otimes 3} \cong & \\ & \left(\left[\mathcal{H}_L^{\otimes 3}\right]_+ \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_+\right) \\ \oplus & \left(\left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2} \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_+\right) \oplus \left(\left[\mathcal{H}_L^{\otimes 3}\right]_+ \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}\right) \\ \oplus & \left(\left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2} \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}\right)\end{aligned}$$

and $\mathbf{P}(\mathcal{S}_n)$ the canonical representation of $\mathcal{S}_3$ on it. Then, $\mathbf{P}(s \in \mathcal{S}_n) = \mathbf{P}_L(s) \otimes \mathbf{P}_R(s)$ implies that: i) the subspace in the first line is in the symmetric subspace of $(\mathbb{C}^4)^{\otimes 3}$, i.e. in $\left[(\mathbb{C}^4)^{\otimes 3}\right]_+$ (and so we do not care about it), ii) the subspace in the second line is in $\left[(\mathbb{C}^4)^{\otimes 3}\right]_{\lambda_2}$ and iii) a nonzero subspace of the subspace in the third line is also in $\left[(\mathbb{C}^4)^{\otimes 3}\right]_{\lambda_2}$. To see this we note that the action of $\mathcal{S}_3$ on $\left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2} \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}$ is non-commutative and since the only irrep of $\mathcal{S}_3$ in which the representation of $\mathcal{S}_3$ is non-commutative is λ_2, therefore by decomposing the action

of $\mathbf{P}(\mathcal{S}_n)$ on $\left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2} \otimes \left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}$ to irreps one should find a λ_2 irrep.

This implies that in $\left[(\mathbb{C}^4)^{\otimes 3}\right]_{\lambda_2}$, $\left[\mathcal{H}_L^{\otimes 3}\right]_{\lambda_2}$ couples to both $\left[\mathcal{H}_R^{\otimes 3}\right]_{+}$ and $\left[\mathcal{H}_R^{\otimes 3}\right]_{\lambda_2}$. This in turn will imply that there is no one-to-one relation between the irreps of U(2) which show up in representations $\mathbf{Q}_L(\mathrm{U}(2)) \otimes I_R$ and $I_L \otimes \mathbf{Q}_R(\mathrm{U}(2))$ in the subspace $\left[(\mathbb{C}^4)^{\otimes 3}\right]_{\lambda_2}$.

Therefore, this example is a counter-example to the above conjecture.

Appendix D: Common figures of merit

Here, we present two examples of common figures of merits and we show that they can be accommodated within our framework.

1. Cost functions

Suppose that $\vec{\mathfrak{s}}(\rho)$ is a vector of parameters to be estimated. As described earlier, any estimation scheme, consisting of a choice of measurement and a post-processing of its outcome, can be described by a POVM $M : \sigma(\Omega) \to \mathrm{End}((\mathbb{C}^d)^{\otimes n})$. In this case, the outcome space Ω must correspond to the range of $\vec{\mathfrak{s}}$. The probability of obtaining an estimate $\vec{s}_{\mathrm{est}}$ of $\vec{\mathfrak{s}}(\rho)$ in the l-dimensional interval $\vec{\Delta} \subseteq \mathbb{R}^l$ is

$$q_M(\vec{s}_{\mathrm{est}} \in \vec{\Delta}|\rho) = \mathrm{tr}(M(\vec{\Delta})\rho). \tag{D1}$$

We will assume that it is possible to express this in terms of a probability *density* function $p_M(\vec{s}_{\mathrm{est}}|\rho)$, defined by

$$q_M(\mathrm{d}\vec{s}_{\mathrm{est}}|\rho) = p_M(\vec{s}_{\mathrm{est}}|\rho)\mathrm{d}\vec{s}_{\mathrm{est}}. \tag{D2}$$

Now suppose that the performance of the estimation scheme will be judged by a cost function (we follow Ref. [5]). The most basic case would be a function of the form $C(\vec{s}_{\mathrm{est}}, \vec{\mathfrak{s}}(\rho))$, which represents the cost of estimating $\vec{s}_{\mathrm{est}}$ when the true value of the parameters is $\vec{\mathfrak{s}}(\rho)$. More generally, however, the cost may depend on more than just the value of the parameters to be estimated. Therefore, some subset of the elements of $\vec{\mathfrak{s}}(\rho)$ might not be included among the parameters to be estimated, being only required to compute the cost, as discussed earlier in this section.

The average cost of the estimation strategy M for the state ρ is

$$\bar{C}(\rho) \equiv \int \mathrm{d}\vec{s}_{\mathrm{est}}\; C(\vec{s}_{\mathrm{est}}, \vec{\mathfrak{s}}(\rho))\; p_M(\vec{s}_{\mathrm{est}}|\rho), \tag{D3}$$

and the expected cost of the estimation strategy M given the prior p is

$$\langle C \rangle \equiv \int d\rho\; p(\rho)\bar{C}(\rho). \tag{D4}$$

The joint probability density for the state being ρ and the estimation strategy M yielding $\vec{s}_{\mathrm{est}}$ is clearly

$$p_M(\vec{s}_{\mathrm{est}}, \rho) = p(\rho)p_M(\vec{s}_{\mathrm{est}}|\rho).$$

Given the definition of the vector of parameters $\vec{\mathfrak{s}}(\rho)$, one can compute from $p_M(\vec{s}_{\mathrm{est}}, \rho)$ the joint probability density for the vector of parameters having value $\vec{s}_0$ and the estimation strategy M yielding $\vec{s}_{\mathrm{est}}$, denoted $p_M(\vec{s}_{\mathrm{est}}, \vec{S} = \vec{s}_0)$. Defining $p(\vec{S} = \vec{s}_0)$ as the marginal on $\vec{S}$ of this joint density, it is always possible to decompose the joint density as

$$p_M(\vec{s}_{\mathrm{est}}, \vec{S} = \vec{s}_0) = p_M(\vec{s}_{\mathrm{est}}|\vec{S} = \vec{s}_0)p(\vec{S} = \vec{s}_0). \tag{D5}$$

It follows that the expected cost can be written as

$$\langle C \rangle \equiv \int d\vec{s}_{\mathrm{est}} \int d\vec{s} p(\vec{S} = \vec{s}) \times C(\vec{s}_{\mathrm{est}}, \vec{s}) p_M(\vec{s}_{\mathrm{est}}|\vec{S} = \vec{s}). \tag{D6}$$

This figure of merit is clearly a functional of $p_M(\vec{s}_{\mathrm{est}}|\vec{S} = \vec{s})$. Because this can be defined in terms of the probability measure $q_M(\mathrm{d}\vec{s}_{\mathrm{est}}|\vec{S} \in \vec{\Delta})$, the figure of merit can be considered a functional of the latter and hence the condition of corollary 15 applies. It follows that if the problem has gauge symmetry $G_{\mathcal{A}}$ and satisfies the assumptions of theorem 13 (or theorem 11), then the optimal estimation can be achieved with POVMs restricted to $\mathcal{A}^{\otimes n}$.

2. Mutual information

The framework for estimation that we have proposed can contend with cases that are more general than those wherein the figure of merit is a cost function. A common figure of merit is the *mutual information* between the actual and guessed values of the parameters to be estimated. Here, we show that this can also be accommodated within our framework.

Note first that the probability density for the actual value being $\vec{S} = \vec{\mathfrak{s}}(\rho) = \vec{s}_0$ and the guessed value being $\vec{s}_{\mathrm{est}}$ given estimation strategy M is defined in Eq. (D5). Denoting the marginal on $\vec{s}_{\mathrm{est}}$ by $p_M(\vec{s}_{\mathrm{est}}) = \int \mathrm{d}\vec{s}\; p_M(\vec{s}_{\mathrm{est}}, \vec{S} = \vec{s})$, the mutual information can be defined as

$$I(\vec{s}_{\mathrm{est}} : \vec{S}) = \int \mathrm{d}\vec{s}\; \mathrm{d}\vec{s}_{\mathrm{est}}\; p_M(\vec{s}_{\mathrm{est}}, \vec{s}) \log \frac{p_M(\vec{s}_{\mathrm{est}}, \vec{s})}{p_M(\vec{s}_{\mathrm{est}})p(\vec{s})}. \tag{D7}$$

Given that $p_M(\vec{s}_{\mathrm{est}}, \vec{s})$ and $p_M(\vec{s}_{\mathrm{est}})$ are functionals of $p_M(\vec{s}_{\mathrm{est}}|\vec{S} = \vec{s})$, the mutual information is a functional of $p_M(\vec{s}_{\mathrm{est}}|\vec{S} = \vec{s})$ as well, so the condition of corollary 15 applies. Once again, It follows that if the problem has gauge symmetry $G_{\mathcal{A}}$ and satisfies the assumptions of theorem 13, then the optimal estimation can be achieved with POVMs restricted to $\mathcal{A}^{\otimes n}$.

[1] R. Goodman and N. R. Wallach. *Representations and Invariants of the Classical Groups* Cambridge University Press, (1998).

[2] A. Harrow, *Applications of coherent classical communication and the Schur transform to quantum information theory*, PhD thesis, MIT, (2005), Arxiv preperint arXiv:quant-ph/0512255.

[3] A. Hayashi, M. Horibe, T. Hashimoto, Phys. Rev. A **73**, 062322 (2006).

[4] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Scuola Normale Superiore (Monographs), (2011).

[5] G. Chiribella, *Optimal estimation of quantum signals in the presence of symmetry*, PhD thesis, University of Pavia, Pavia, Italy (2006).

[6] I. Marvian and R. W. Spekkens, Arxiv preprint arXiv:quant-ph/1105.1816 (2011).

[7] I. Marvian and R. W. Spekkens, Under preparation.

[8] P. Zanardi and M. Rasetti, Phys. Rev. Lett. **79**, 3306 (1997); P. Zanardi, Phys. Rev. A **63**, 012301 (2000).

[9] E. Knill *et al.*, Phys. Rev. Lett. **84**, 2525 (2000); J. Kempe *et al.*, Phys. Rev. A 63, 042307 (2001).

[10] S. D. Bartlett, T. Rudolph and R. W. Spekkens, Phys. Rev. Lett.**91**, 027901 (2003).

[11] C. W. Helstrom, *Quantum detection and estimation theory*, Academic press (1976).

[12] M. Keyl and R. F. Werner, Phys. Rev. A 64, 052311 (2001).